THE Informal Labour Market in Brazil: Job Queue, Trade Liberalisation and Minimum Wage

Fabio Veras Soares

April, 2005

A Thesis submitted for the Degree of Doctor in Philosophy to University College London, The University of London.

Abstract

The main concern of this thesis is to investigate several aspects of the informal sector in Brazil. First, it aims to establish and analyze in depth some stylized facts such as the increase in the size of the informal sector during the 1980's and the 1990's and the fall in the wage gap between formal and informal workers in the mid-1990's. Second, it tests whether informal sector workers queue for formal jobs. Third, it analyses the effect of policy changes on the informal sector. In this regard, it investigates the impact of the trade liberalisation process of the early 1990's on the proportion of informal workers and on the wage gap between formal and informal workers, and the impact of minimum wage hikes on the transitions to and from both the formal and informal sectors.

As for the main results we find that: 1) Informal sector workers queue for formal jobs and that non-white, female, less educated, "new entrants" and former informal workers are the groups with the lower probability of being chosen from the queue for formal jobs; 2) the fall in the wage gap between registered and non-registered workers in the manufacturing sector was affected by trade-related variables, particularly, by the import penetration ratio. However, we do not find robust evidence that trade liberalization had a substantial impact on the fall in the proportion of registered Workers; 3) there were no disemployment effects of the minimum wage in the 1980's, but there were some disemployment effects in the 1990's for both formal and informal sector workers. However, we find no strong evidence that minimum wage hikes led to transitions from formal to informal sector or to self-employment either in the 1980's or in the 1990's.

Acknowledgments

I would like to thank CAPES (Brasilia, Brazil) and IPEA for financing my PhD studies. My supervisors Stephen Machin and Donald Verry for their comments, suggestions, and support. Ricardo Paes de Barros, Maria Cristina Cacciamali, Francisco Carneiro and Jorge Arbache for their support and incentive for my PhD. Jorjao, in special, was a (e-mail) present and wise adviser throughout this period. The Brazilians at UCL: Sara, Miguel and Kiko (Carlos Henrique Corseuil) for insightful discussions and help with daily issues. Special thanks to Kiko who while in Brazil at IPEA helped me a lot in gathering the data and elucidating doubts. Sergei Soares, Luciana Servo, Marcela Aurelio, Honorio Kume, Carmem Falcao, Wilson Morgado, Aguinaldo Maciente and Eustaquio Reis were also of great help at IPEA; Renato Fonseca at CNI, Marileni at IBGE , all helped and guided me through the data and their devious details.

I also would like to acknowledge the contributions of the participants of my transfer seminar at UCL, specially Christian Dustman and Nicholas Rau. The participants of the VI IZA Summer School in Munich, specially, Jeff Smith, Ana Rute Cardoso, Jean Ries, and Habiba Djebari for their comments and suggestions on the paper that I presented there and on which the Chapter 5 of this thesis is based. Chiara Criscuolo, Ekin Birol, Emma Aguila and Manuela Angelucci were magnificent friends at UCL who helped me a lot academically and psychologically in the past four years, without these four good souls my life during the PhD would have been much tougher. A special thanks to all academics who I have contacted to clarify details of their papers, to ask copies of unpublished papers, and another thank you to all participants of several software's discussion lists who saved me in the most desperate moments when my programming skills completely failed me. For all these anonymous, some of them not so anonymous, and occasional helpers my most sincere thank you.

I would like to thank my friends in the little Wivenhoe for their support and encouragement: Cinara Nahra, Seth Hague, Cris Berta, Elena Bardasi, Cheti Nicoletti, Mada and Glaucia Falsarelli. My old friends who gave their support through e-mail: Edvaldo (Didi), Anna Ozorio de Almeida and Pedro Crosetti and my family in Brazil. Finally, I would like to thank Marcelo for his patience, help and love during these four years.

Declaration

No Part of the thesis has been presented to any university for any degree

Contents

1 Introduction **8**

1.1 Motivation 8

1.2 Layout of the Thesis 11

2 Survey of the Literature and Some Stylised Facts of the Informal Sector in Brazil **14**

2.1 The Literature on Informal Sector in Developing Countries 14

2.2 Formal and Informal Sector in Brazil: Some Stylized Facts 20

2.2.1 Data 21

2.2.2 Some Descriptive Statistics 21

2.2.3 Changes in Employment Structure during the 1980's and the 1990's 24

2.2.4 Evolution of Wage Differential Between Registered and Non-registered Workers 27

2.2.5 Returns to Education for Registered and Non-Registered Workers 33

2.2.6 Decomposing Wage Differential and Accounting Inequality for Registered and Non-registered Workers: 1981-1999 36

2.3 Concluding Remarks 48

3 Do Informal Workers Queue for Formal Jobs in Brazil? **50**

3.1 Introduction 50

3.2 Endogenous Switching Regression in the Presence of Job Queue . . . 54

3.3 Estimation Procedure 57

3.4 Data 66

3.5 Results 69

3.5.1 Univariate Probit Reduced Form 69

3.5.2 Bivariate Probit Reduced Form 71

3.5.3 Job Queue Tests and the Length of the Queue 77

3.5.4 Wage Equations for Formal and Informal Workers 79
3.5.5 The Role of Wage Differential on Sector Allocation 86
3.6 Robustness Checks 91
3.6.1 Tenure as an Exclusion Restriction in the Bivariate Probit Models 91
3.6.2 The Stability of the Parameters of the Wage Equations 96
3.7 Conclusion 102

4 The Impact of Trade Liberalisation on the Informal Sector in Brazil 105
4.1 Introduction 105
4.2 Registered and Non-registered Workers in the Manufacturing Sector in Brazil 107
4.3 Trade Liberalisation and Labour market 109
4.3.1 A Brief Overview of the Literature on the Impact of Trade Liberalisation on Developing Country Labour Markets. 109
4.3.2 Trade Liberalisation and Segmented Labour Market. 115
4.4 Some Features of the Trade Liberalisation Process in Brazil 119
4.5 Empirical Strategies 126
4.5.1 The Impact of Trade Liberalisation on the Manufacturing Sector. 127
4.5.2 Spillover Effects of Trade Liberalisation 136
4.5.3 Pseudo-cohort Approach 141
4.6 Conclusion 150

5 Minimum wage and the Informal Sector in Brazil 152
5.1 Introduction 152
5.2 The Recent Literature on the Disemployment Effects of the Minimum Wage 154
5.2.1 The International Evidence 154
5.2.2 Evidence for Brazil 156
5.3 Minimum Wages in Brazil 158
5.3.1 The Minimum Wage Policy between 1982 and 1999 158
5.3.2 A Tale of Two Decades and Two Sectors 161
5.4 Methodology 166
5.4.1 Difference-in-Differences with Pseudo-Experiment 166
5.4.2 Choosing control groups 168
5.4.3 The Common Support Problem and Matching Methods 173
5.5 Data 178

5.6 Results 180
5.6.1 Minimum Wage with Indexation: The period 1982 - 1985 . . . 180
5.6.2 Minimum wage without indexation: the period 1995-1999 . . . 189
5.6.3 Assessing Comparison Groups for the Non-parametric Results 197
5.7 Conclusion 201

6 Conclusion **205**

7 References **212**

A Appendix - Chapter 2 **222**

B Appendix - Chapter 4 **226**

C Appendix - Chapter 5 **229**

List of Figures

2.1 Proportion of Registered and Non-Registered Workers (in %) - 1981 - 1999 25
2.2 Wage Premium for Registered Workers - 1981 - 1999 26
2.3 Log Hourly Real Wage Rate - 1981 - 1999 29
2.4 Wage Differential between Registered X Non-registered (Filter based) - 1981 - 1999 29
2.5 Wage Differential between Registered X Non-registered (Survey based) - 1989 - 1999 30
2.6 Wage Differential between Registered X Non-registered (Survey based and more controls) - 1989 - 1999 31
2.7 Wage differential between Male X Female (Filter based) - 1981 - 1999 31
2.8 Wage differential between Male X Female (Survey based) - 1989 - 1999 32
2.9 Wage differential between Male X Female (Survey based and more controls) - 1989 - 1999 32
2.10 Relative Returns to Education (Full Sample) - 1981 - 1999 34
2.11 Relative Returns to Education (Registered Sample) - 1981 - 1999 . . 34
2.12 Relative Returns to Education (Non-Registered Sample) - 1981 - 1999 35
2.13 Standard Deviation of the Log Hourly Real Wage - 1981 - 1999 41
2.14 Standard Deviation of the Residuals - 1981 - 1999 41

4.1 Proportion of Registered Workers in the Manufacturing Sector - 1981 - 1999 108
4.2 Wage Differential between Registered and Non-registered Workers in the Manufacturing Sector (in %) - 1981 - 1999 109
4.3 Wage Premium and Trade Measures (Pooled Sample) - 1981 - 1999 . 124
4.4 Proportion of Registered Workers and Trade Measures (Pooled Sample) - 1981 - 1999 125
4.5 Nominal Tariff and Trade Measures (Long Changes: 1987-1998) . . . 130
4.6 Wage Premium and Trade Measures (Long Changes: 1987-1998) . . . 130

4.7 Proportion of Registered Workers and Trade Measures (Long Changes: 1987-1998) 131
4.8 Proportion of Registered workers by cohort (1981 - 1999) 142

5.1 Real Minimum Wage in Six Metropolitan Areas - 1982 - 1999 160
5.2 Kernel Density of Log Wages: 1985 163
5.3 Kernel Density of Log Wages: 1995 164
5.4 Wage Change by Percentiles due to Minimum Wage Increase: 05/1985 and 05/1995 165
5.5 Wage Distribution and Treated and Control Groups: 04/1995, 05/1995 and 09/1995 167
5.6 Job loss before and after minimum wage increases in 1985 and 1995 . 172

A.1 Relative Returns to Education (Survey based: Full Sample) - 1989 - 1999 223
A.2 Relative Returns to Education (Survey based: Registered Sample) - 1989 - 1999 223
A.3 Relative Returns to Education (Survey based: Non-Registered Sample) - 1989 - 1999 224
A.4 Relative Returns to Education (Survey based and more controls: Full Sample) - 1989 - 1999 224
A.5 Relative Returns to Education (Survey based and more controls: Registered Sample) - 1989 - 1999 225
A.6 Relative Returns to Education (Survey based and more controls: Non-Registered Sample) - 1989 - 1999 225

List of Tables

2.1 Summary of Studies on Formal x Informal Wage Differential and Segmentation . . . 18
2.2 Descriptive Statistics for Non-Registered and Registered Workers in Selected Years (1981, 1990 and 1999) . . . 23
2.3 Oaxaca-Blinder Average Wage Decomposition: 81-99 . . . 38
2.4 Oaxaca-Blinder Wage Gap Decomposition: 81-99 . . . 38
2.5 JMP Decomposition of Changes in the Wage Gap between Formal and Informal Workers: 1981-1999 . . . 40
2.6 The Effect of the Decline in Wage Differential on the Variance of Log Hourly Wage . . . 42
2.7 Factor Contribution to Wage Inequality and to Change in Inequality: 1981 - 1999 . . . 44
2.8 Juhn-Murphy-Pierce Changes in Inequality Decomposition . . . 47

3.1 Descriptive Statistics for 1990: Full Sample, Registered and Non-registered Workers) . . . 68
3.2 Univariate Probit Specification . . . 70
3.3 Bivariate Probit Specifications . . . 72
3.4 Descriptive Statistics for Informal Sector Workers: 1990 . . . 76
3.5 LR Test for Universal Queue and No Queue Hypotheses . . . 78
3.6 Probabilites of a) Being Chosen From the Queue Conditional on Being in the Queue (PCF), b) Being in the Queue (PIQ), c) Being in the Formal Sector (PF); and Length of the Queue (Q). . . . 79
3.7 Wage Equations: Univariate-based corrections . . . 82
3.8 Wage Equations: Bivariate-based Corrections . . . 84
3.9 Effect of the correlation between Unobservables in the Choice Equations and in the Outcome Equation . . . 86
3.10 Structural Univariate Probit . . . 87
3.11 Abowd-Farber Bivariate Probit Specification . . . 89

3.12 Bivariate Probit with Sample Selection Structural Bivariate Probit Specification . . . 91
3.13 Sensitivity of the Abowd-Faber Bivariate Probit to the Inclusion of Tenure in the CFQ Equation . . . 92
3.14 Sensitivity of the Poirier Bivariate Probit to the inclusion of Tenure in the CFQ . . . 93
3.15 Sensitivity of the Bivariate Probit with Sample Selection to the Inclusion of Tenure in the CFQ equation . . . 94
3.16 OLS Wage Equations Based on the Univariate Probit with Different Exclusion Restrictions . . . 97
3.17 MLE Wage Equations Based on the Univariate Probit with Different Exclusion Restrictions . . . 98
3.18 Wage Equations Based on the Abowd-Farber Bivariate Probit with Different Exclusion Restrictions . . . 99
3.19 Wage Equations based on the Bivariate probit with Sample Selection with Different Exclusion Restrictions -1st part (cont.) . . . 100
3.20 Wage Equations based on the Bivariate probit with Sample Selection with Different Exclusion Restrictions-2nd part . . . 101

4.1 Nominal Tariff and Import Penetration Coefficient (in %) . . . 121
4.2 Spearman Rank Cross-correlation in 1987 . . . 122
4.3 Spearman Rank Cross-correlation in 1998 . . . 123
4.4 Spearman Rank Correlations between 1987 and 1998 . . . 124
4.5 Wage Premium: Fixed Effect with Time Dummies (Contemporaneous Regressors) . . . 133
4.6 Proportion of Registered Workers: Fixed Effect with Time Dummies (Contemporaneous Regressors) . . . 134
4.7 Wage Premium: Fixed Effect with Time Dummies (Lagged Regressors) 135
4.8 Proportion of Registered Workers: Fixed Effect with Time Dummies (Lagged Regressors) . . . 136
4.9 Regional Wage Premium: Fixed Effect with Time Dummies (Contemporaneous Regressors) . . . 138
4.10 Regional Proportion of Registered Workers: Fixed Effect with Time Dummies (Contemporaneous Regressors) . . . 139
4.11 Regional Wage Premium: Fixed Effect with Time Dummies (Lagged Regressors) . . . 140

4.12 Regional Proportion of Registered Workers: Fixed Effect with Time Dummies (Lagged Regressors) 141
4.13 Number of Years Observed for Each Industry-cohort Cell 144
4.14 Pseudo-Cohort: Fixed Effects with Time Dummies for log Hourly Wage 147
4.15 Pseudo-Cohort: Fixed-Effect with Time Dummies for the Proportion of Registered Workers 149

5.1 Monthly Transitions Between Employment Status: 1982-1986 181
5.2 Monthly Transitions Between Employment Status: 1995-1999 182
5.3 Nominal Minimum Wage and Nominal Increases 183
5.4 Changes in Transition to Non-employment: Probit (1982-1985) 184
5.5 Changes in Transition to Informality: Probit (1982-1985) 185
5.6 Changes in Transition to Non-Employment: Kernel Propensity Score Matching (1982-1985) 187
5.7 Changes in Transition to Informality: Kernel Propensity Score Matching (1982-1985) 187
5.8 Changes in Transition to Non-employment: Nearest Neighbour Score Matching (1982 - 1985) 189
5.9 Changes in Transition to Informality: Nearest Neighbour Score Matching (1982-1985) 189
5.10 Changes in Transition to Non-employment: Probit (1995-1999) 191
5.11 Changes in Transition to Informality: Probit (1995-1999) 193
5.12 Changes in Transition to Non-Employment: Kernel Propensity Score Matching (1995-1999) 194
5.13 Changes in Transition to Informality: Kernel Propensity Score Matching (1995-1999) 195
5.14 Changes in Transition to Non-Employment: Nearest Neighbour Propensity Score Matching (1995-1999) 196
5.15 Changes in Transition to Informality: Nearest Neighbour Propensity Score Matching (1995-1999) 197
5.16 Difference in Probability Before and After the Treatment for the Control Group (Pooled Sample) 198
5.17 Difference in Probability Before and After the Treatment for Control Group (Formal Sector) 200
5.18 Difference in Probability Before and After the Treatment for Control Group (Informal Sector) 202
5.19 Summary of the Effect of the Minimum Wage Hikes 203

B.1 Effective Tariff in % . . . 226
B.2 Nominal Tariff in % . . . 227
B.3 Import Penetration Ratio in % . . . 227
B.4 Export Orientation Ratio in % . . . 227
B.5 Dummy Coefficient for Registered Workers . . . 228
B.6 Proportion of Registered Workers . . . 228

C.1 Changes in Transition to Non-employment: Probit (Excluding 1.5, 2 and 3 minimum wages) . . . 230
C.2 Changes in Transition to Non-employment: Probit (conditional on Wages two Months before the Minimum Wage Increase) . . . 231
C.3 Changes in Transition to Informality: Probit (Robustness Check - Formal Sector) . . . 232

Chapter 1

Introduction

1.1 Motivation

Brazil had 36 million private sector wage workers above the age of 10 years in 1999[1]. Out of that, 14 million were in the informal sector, i.e, their job contract was not registered in their work card. The main concern of this thesis is to investigate several aspects of the informal (non-registered) sector in Brazil. First, it aims to establish and analyze in depth some stylized facts such as the increase in the size of the informal sector during the 1980's and the 1990's and the fall in the wage gap between formal and informal workers in the mid-1990's. Second, it tests whether informal workers queue for formal jobs. Third, it analyses the effect of policy changes on the informal sector. This is an issue that has not received much attention in the literature. In this regard, it investigates the impact of the trade liberalisation process of the early 1990's on the proportion of informal workers and on the wage gap between formal and informal workers, and the impact of minimum wage hikes on the transitions from formal and informal sectors to disemployment and from formal to the informal sector.

The size of the informal sector in Brazil - almost 40% of wage workers - is in itself something that demands an explanation. This figure means that a sizable portion of workers are not entitled to benefits such as unemployment insurance and do not contribute to social security. This is both a social and a fiscal problem in a country that has been struggling to replicate the high growth rates witnessed until the 1970's. The challenge is to revive growth without falling back in the

[1]This is the aggregate figure for the whole country (except the rural North region) and excludes public sector wage workers (4.9 million), self-employed workers (16.8 million) and domestic workers (5.9 million). This data is published by the National Statistics Office, IBGE - Instituto Brasileiro de Geografia e Estatisticas. Website: http://www.ibge.gov.br/.

hyper-inflationary process that made it impossible to grow in a consistent basis from the early 1980's until middle 1990's, the so-called *lost decade*. The proportion of the informal sector increased 10% during these 15 years. It did so even in the manufacturing sector in which it used to be relatively unimportant[2]. Paradoxically, the wage gap between formal and informal workers also decreased between 1981 and 1999. Most of the reduction in the wage gap occurred after the market-oriented reforms (e.g. privatization, trade liberalisation, deregulation) of the early 1990's and after the enactment of the new Federal Constitution of 1988[3]. It is true that there were some episodical reductions in the wage gap (e.g. 1986 and 1990), but it was only from 1992 onwards that this reduction was not significantly reversed by changes in the business cycle or by the melting down of price and wage controls of unorthodox stabilization plans. The reduction in the wage gap between formal and informal workers helped to reduce slightly wage inequality, but the latter is still extremely high when compared to other similar countries[4]. Furthermore, the wage package of formal workers contains, besides the mandatory benefit associated with the registration, several fringe benefits that are not readily accessible to non-registered workers (e.g. transport and food vouchers), so that the actual inequality is likely to be higher than the one reported in the estimates. The productive attributes of non-registered workers improved over this period, and this fact can explain part of the wage gap reduction, but not all of it. Another interesting change that occurred during this period was the fact that non-registered workers became over-represented among the minimum wage earners. In addition to that, there is also some evidence that their wage increases were linked to minimum wage hikes.

The changes and stylised facts highlighted above raise some questions that we will try to answer in this thesis. Is it the case that 40% of wage workers choose to join the informal sector because they have comparative advantage in that sector? If so, how do we explain that when asked whether or not they would accept a formal job offer, 70% of the informal workers say "yes"?[5] A switching regression model

[2]Many commentators argue that the fall in the proportion of workers in manufacturing industry and the increase in the proportion of workers in the service industry are the main culprit for the increase in the proportion of informal workers. However, as shown in Chapter 2 and 3 of this Thesis and in Ramos (2002), the proportion of informal workers increased within the manufacturing sector, so that changes in the sectoral structure of employment cannot be entirely responsible for this phenomenon.

[3]The New Constitution created several new rights to formal workers and reduced the maximum workweek.

[4]According to the 2002 World Development Report published by the World Bank, Brazil Gini's index of 0.61 is among the highest in the world, comparable to Central African Republic (0.62), Sierra Leone (0.63) and Nicaragua (0.61), and well above Argentina (0.45) and Mexico (0.51).

[5]See Pero and Urani (1994) and Chapter 3 of this thesis.

that takes into account the two decision process that determine the formal status of workers should be modelled in order to have a clear picture of the sector allocation process. Hence, the decision to join the queue for formal sector job and the decision of hiring a worker from that queue and the correlation between the unobservable of these two processes could offer a better understanding on which type of workers profit from getting a formal job.

As for the increase in the size of the informal sector and the fall in the wage gap between formal and informal workers, one could ask whether trade liberalisation had any causal relationship with there phenomena given the timing of those changes? Did the trade liberalisation process curb the rents of the formal sector workers in the protected sectors? Did such process have a spillover effect on the rest of the economy? If not, did the tradable sector experienced a fall in the wage premium of registered workers more pronounced than the non-tradable sector?

Another hypothesis usually raised in the literature to explain the surge and expansion of the informal sector is the existence of a binding minimum wage. Were the increases in the minimum wages observed after the *Plano Real* responsible for the displacement of workers from the formal sector? Was the indexation of informal sector wages to minimum wage hikes responsible for the compression in the wage gap between formal and informal sectors?

At this point it is necessary to clarify the concept of informal sector we will be using. Throughout this thesis we will be referring to the informal sector as the set of workers whose contract is not registered in his/her work-card (*carteira de trabalho*). According to the Brazilian legislation, registered workers are the ones whose labour contract is registered on their work-card. This registration entitles them to several wage and non-wage benefits such as 30 days of paid holiday per year, contribution for social security, right to request unemployment benefit in case of dismissal, monetary compensation if dismissed without a fair cause, maternity and paternity paid leave and so on. Differently, non-registered workers have informal contracts, which are illegal and not registered in their work-card; in general any benefit such as paid holiday must be agreed with the employer on a case-by-case basis. Moreover, non-registered workers do not have access to any of the government-administered benefits related to the labour market, such as unemployment benefit and severance payment.

It is important not to confound this classification with the *ILO* or *ILO*-related definitions of the informal sector. In general, these definitions comprise non-professional self-employed, employers and employees in small firms with cut points varying from 5 to 15 employees and non-paid workers [Maloney (1997), Gong et al. (2000)]. Our classification of registered and non-registered workers is an institutional one, in which employers avoid some sort of regulation, in this case, compliance with the labour code. Other possible definitions in the institutional framework are: the lack of contribution for social security as in Verry and Araujo (1996) or working in the underground economy. We prefer the registered/non-registered classification because it allows us to concentrate on the labour market strictly defined, i.e., on employees who work for a firm and receive monetary payment.

Another necessary clarification is the exclusion of self-employment of the analysis. This group of workers has also increased over time, but despite recognizing that ideally the process of occupational choice should take this sector into account, we decide not to analyse it for three reasons. First, factors such as managerial ability and entrepreneurial talent play a crucial role in the sector allocation decision of self-employed and small-firms owners (Yamada, 1996), therefore, focusing on a sample of employees should reduce possible selectivity problems[6]. Second, it is widely recognised the difficulties in comparing wages of employees and earnings from self-employed and employers that, in general, contain more than their net remuneration, including also the remuneration of capital. Third, it would be hard to separate out the typical self-employed from liberal professionals on the data used on this thesis.

1.2 Layout of the Thesis

This thesis is structured as follows. In Chapter 2 we review the literature on the informal sector in developing countries and analyse some stylised facts on the informal sector in Brazil. The first part of this chapter describes the different approaches to the informal sector and how the concept of an informal sector has been dealt with theoretically. It highlights the role of efficiency-wage models in its shirking version and firm heterogeneity in order to explain the wage differential between formal and informal workers. It also reviews the empirical literature for developing countries with emphasis on the way selectivity issues were handled when assessing the wage differential between formal and informal workers. The second part of the chapter deals with the stylized facts that we have mentioned in this introduction. It shows the evolution of the informal sector during the 1980's and the 1990's, it analyses the changes in the employment structure and it investigates the determinants of the changes in the wage differential and in the wage inequality for all employees and for the subsample of formal and informal workers, separately. Special emphasis is given

[6]Even in Chapter 3 when we deal with selectivity issues the inclusion of an additional sector would make the estimation procedure much more complicated, especially in a context of a bivariate probit with sample selection

to the role of the fall in the wage gap in curbing wage inequality.

In Chapter 3 we investigate the existence of a job queue for formal (registered) jobs in an endogenous switching regression framework. This approach aims at correctly specifying the allocation process in the presence of queuing and getting unbiased estimates of the wage equation for both formal and informal sectors in order to evaluate the role of wage differential in determining sector allocation. We estimate three types of bivariate probit specifications in order to evaluate the sensitivity of the results to different assumptions about the sector allocation process. In particular, we assess the sensitivity of the job queue estimates to the usual assumption of partial observability using subjective questions on the desire of informal (non-registered) workers to switch to a formal job. We find strong evidence that informal workers queue for formal jobs and that non-white, female, less educated, "new entrants" and former informal workers are the groups with the lower probability of being chosen from the queue for formal jobs conditional on being in the queue. We also find that wage differential has a major role in the decision to join the queue. Robustness checks on the bivariate specifications indicate that the bivariate probit with sample selection, that exploits the subjective question about the desire to move to the informal sector, is much less sensitive to changes in specification.

In Chapter 4, we assess whether or not the trade liberalisation process in Brazil had any effect both on the reduction in the wage differential between registered and non-registered workers and on the increase in the proportion of non-registered workers. We discuss the channels through which trade liberalisation could affect these two variables and put forward three empirical approaches to test the existence of any correlation between them. Our results suggest that the fall in the wage gap between registered and non-registered workers within the manufacturing sector was affected by trade-related variables, particularly by the import penetration ratio. However, we do not find robust evidence that trade liberalisation had a substantial effect on the fall in the proportion of registered workers.

In Chapter 5, we investigate the effect of the minimum wage on employment transitions for both formal and informal sector workers. We estimate the probability of becoming non-employed (unemployed or out of the labour force) and the probability of moving to the informal sector[7] after minimum wage hikes. We estimate these effects separately for periods with high and low inflation to assess how agents react to minimum wage hikes under different inflationary expectations, particularly under different degrees of wage indexation in the formal and informal sectors. Workers affected by minimum wage increases are compared with similar workers further up

[7]In this latter case, we investigate the transition only for formal sector workers.

the wage distribution. In order to control for heterogeneity between the treated minimum wage workers and the control groups we use a difference-in-difference approach that compares treated and control groups in periods with a nominal increase in the minimum wage with periods with no increase. In this last case the control and treated groups are defined as if there had been an increase in the minimum wage (pseudo-experiment). This strategy is applied in a parametric way via probit estimates and also in a nonparametric way using different propensity score matching methods. We also experiment with different control groups due to the possible spill-over effect (mainly for multiple of minimum wage earners) triggered by the strong wage indexation observed during the 1980's. Our results suggest that there were no disemployment effects of the minimum wage in the 1980's, but there were some disemployment effects in the 1990's for both formal and informal sector workers. However, we find no strong evidence that minimum wage hikes led to transitions from the formal to the informal sector or to self-employment either in the 1980's or in the 1990's.

Chapter 6 summarizes the results of the previous chapters and discusses some of the implications of results for public policy and ideas for future research on these topics.

Chapter 2

Survey of the Literature and Some Stylised Facts of the Informal Sector in Brazil

2.1 The Literature on Informal Sector in Developing Countries

As in most developing countries, the Brazilian labour market is characterized by sharp differences in the way its citizens are linked to it. This contrast is usually called "segmentation" or "dualism" and may refer among other factors to differences between the modern formal sector and the traditional urban sector, to differences between small and large firms, and to the wage differential between workers in the formal and in the informal sectors.

The first typology derives directly from the work of Fields (1975) – in the tradition of the development economics field – and treats the urban traditional sector or the "murky" sector as a buffer for unemployed workers who migrate from rural areas attracted by job opportunities in the urban formal labour market. The low unemployment rate observed in developing countries would be due to the fact that workers would stay in the "murky sector" while looking for job in the formal sector (queuing for it). The "murky sector" is comprised, according to this view, by small business that employ low skilled workers since formal employers prefer more educated workers.

The second approach challenges Fields's view of the informal sector as a "waiting stage" and sees the informal sector as the lower end of the distribution of firms in developing countries. This view was born by the *ILO* (1972) report on Kenya and it rejects the idea of the urban informal sector as a "waiting stage" to access a "good" formal job. It regards the informal sector as a permanent source of employment and income. This small-scale or technological-based definition of the informal sector led to the definition of informal sector for purposes of quantification as being comprised of self-employed workers, and employers, employees and non-remunerated workers working in firms with less than a determined employment threshold that, in general, varies from 5 to 15.

The third approach is concerned about the wage differential between workers with similar productive characteristics but allocated in different sectors. Actually, this approach can be traced back to the literature on dual labour market in developed economies that tried to provide evidence that similar workers in different sectors are paid differently, so that the labour market could not be characterized as perfectly competitive. This theory, whose origin dates back to the work of Doeringer and Piore (1971), was afterwards embraced by the different theories of efficiency wages, so that their concept of primary and secondary labour market could be laid on a solid microeconomic basis. Of course, several explanations for the wage differential between individuals with similar productive characteristics are possible within the competitive framework such as compensating differentials, but the bulk of empirical evidence is not consistent with such hypotheses[1]. However, most of these earlier findings were challenged because they did not take into account unobserved heterogeneity among workers that might lead them to be more productive in one sector than in the other. Heckman and Sedlacek (1985), Heckman and Hotz (1986) and Magnac (1991) argue that the sector choice of a worker is based on his/her comparative advantage. A worker chooses to work in the sector in which he/she is more productive and hence where he/she is able to command a higher wage. This allocation process affects the comparison between wage equations from different sectors, since workers found in each sector are not randomly drawn from the population. Former empirical studies based on the comparison of two (or more) different wage equations that do not take into account the allocation process were plagued with selectivity bias[2].

The three approaches presented above are not mutually exclusive. Several models blend elements of these three approaches to explain how the formal-informal sector dichotomy arises. Esfahani and Salehii-Isfahani (1989) build an efficiency wage model in its shirking version that encompasses characteristics of the three approaches

[1]See for instance Krueger and Summers (1987) and the chapter 5 of Saint-Paul (1997) for a review of these studies.

[2]See Dickens and Lang (1985) for an attempt to overcome this criticism and Heckman and Hotz (1986) for a critique of that attempt.

described above. In their model, technological dualism is related to the dichotomy between formal (large firms) and informal (small firms) and it leads to differences in the observability of effort by employers in different sectors. It is assumed that effort is less observable in the formal than in the informal sector, so that workers in small (informal) firms who are perfectly monitored are paid competitive wages, whereas workers in large (formal) firms who are not perfectly monitored are paid efficiency wages. Thus segmentation is generated due to both the cost of monitoring workers and technological dualism. Nevertheless, the wage differential between formal and informal sectors in this model is reduced to a "size" effect on wages. Rauch (1991) builds a model where the formal-informal sector dualism in the labour market is integrated with size dualism via the hypothesis that the minimum wage is only enforced on firms larger than a certain size. The size gap between formal and informal sector firms varies with the wage differential between formal and informal sector workers, which increases with hikes in the minimum wage. In this same vein, Fortin et al. (1997) build a model where formal-informal sector dualism arises endogenously due to firms' heterogeneity and to the assumption that the marginal cost of tax and regulation evasion (e.g. not paying the minimum wage or other mandatory contributions) increases with the size of the firm. This model is also compatible with discontinuity in the size distribution of firms (the so-called missing middle) and with "waiting unemployment" as in Field's model.

These models are interesting in the sense that they offer a benchmark to understand the relationship between several characteristics of the informal sector and how changes in policies can affect the size of the informal sector and the wage differential between workers. However, the empirical literature has emphasized the wage differential studies and we are not aware of many papers that try to assess the effect of changes in policies[3] (e.g. labour law reform, trade liberalisation and minimum wage hikes) on the size of the informal sector or on the wage differential between formal-informal sector workers.

In Table 2.1, we summarize some of the results of selected papers, emphasizing the way they define the informal sector, the methodology used and the key findings. Heckman and Hotz (1986) investigate the hypothesis of segmentation in Panama. Assessing whether low-wage workers have a lower return to education than high wage workers, as predicted by the segmentation hypothesis, they found the opposite result, i.e, if there is any segmentation in the Panamanian labour market, it would

[3]Fortin et al. (1997) build in their paper a computable general equilibrium (CGE) model for Cameroon in order to simulate the effect of several reforms of the tax and regulation system on the informal sector. They find that an increase in the tax rate on profits, in the payroll tax, and in the government set wage rate increase the size of the informal sector.

favour low-wage workers. Even after re-estimating the model with the correction for the "sector-allocation choice", the results still indicate the presence of this "reverse" segmentation. Nevertheless, Heckman and Hotz were not convinced that the available tests would be enough to characterize the presence of segmentation. Among the reasons for this disbelief is the possibility of mispecification of the wage equation. The mispecification would lead to the rejection of equality of the parameters of the wage equations for the two groups even in the absence of segmentation. The main effect of Heckman's work in this area was to lead most of the subsequent research to adjust the estimation of the wage equation to tackle selectivity issues. For this reason most of the studies that investigate the hypothesis of segmentation against the hypothesis of comparative advantage in the sector allocation choice concentrate their analysis on sign of the Inverse Mills Ratio or on the correlation (ρ) between unobservables in the wage equation and unobservables in the "choice" equation. A positive correlation between the unobservables of the two equations would imply that workers selected themselves into formal and informal sector according to their comparative advantages.

The lack of pattern in the way the informal sector is defined across studies makes it difficult to compare the findings reported in Table 2.1. Some studies are worried basically about self-employment, others about unregulated (non-registered) workers, and others follow the ILO definition or some modification of that. Most studies use 2-step Heckman selection and try to correct the wage equation in order to properly evaluate the wage differential. However, not many go beyond the second step to investigate the role of wage differential in a structural framework as in a switching regression model[4]. In general, they only report the reduced form 'choice equation'[5] and the wage equations corrected for selectivity.

One interesting point is that some studies try to disentangle the hypothesis of comparative advantages from the hypothesis of managerial abilities. The idea is that self-employed workers should be positively self-selected because they have a special talent for business, while the same would not necessarily hold for wage workers regardless of their registered status (Yamada, 1996). Comparative advantage hypothesis would prevail only if we could not reject positive selection for all types of workers involved in the estimation process.

In the case of Brazil, very few studies have tried to control for selectivity bias while estimating the wage differential between formal and informal workers. Carneiro

[4]See Chapter 3 for a more complete discussion of this point.

[5]The reduce form 'choice equation' can be estimated as a probit, a logit or even a multinominal logit in the case of more than two sectors, e.g, Yamada (1996).

Study	Country	Definition	Methodology	Key Findings
Blau (1985)	Malaysia	Self-employed	2-step Heckman selection Model	Negative selection in rural areas and positive selection in urban areas.
Heckman and Hotz (1986)	Panama	Low income workers	2-step Heckman selection Model	Reverse segmentation.
Gindling (1991)	Costa Rica	Unprotected workers (wage employees and self-employed)	2-step Heckman selection Model (multinomial logit)	There is segmentation: employer's decision.
Barros et al. (1992)	Brazil	Non-registered and self-employed workers (separately)	Mobility Analysis	Movers to informal sector lose, movers to formal sector gain.
Yamada (1996)	Peru	Self-employed	3-step Heckman selection Model (endogenous switching regression model)	Positive selection for self-employed (managerial ability) and negative selection for non-registered wage workers.
Funkhouser (1997)	Guatemala	Workers in small firms (5 employees or less)	2-step Heckman selection Model	Positive selection for both formal and informal workers .
Marcoullier et al. (1997)	Peru, Mexico and El Salvador	Workers in small firms (4 employees or less) and non-professional self-employed.	2-step Heckman selection Model	Salvadorans and Mexican men: positive selection into the informal sector. No selection in Peru. Negative selection for Mexican women.
Maloney (1999)	Mexico	Workers and owners of small firms (16 employees or less)	Mobility analysis	No segmentation. Wage differential is not a good test for segmentation.
Saavedra and Chong (1999)	Peru	Non-registered and informal self-employed (separately)	2-step Heckman selection Model	Positive selection of non-registered workers.
Gong et al. (2000)	Mexico	Workers in small firms (16 employees or less)	Mobility Analysis	Movers to informal sector lose, movers to formal sector gain.
Carneiro and Henley (2001)	Brazil	Non-registered workers	3-step Heckman selection Model (endogenous switching regression model)	Positive selection for non-registered workers and negative selection for registered workers.
Gong and Van Soest (2002)	Mexico	Piece-workers and self-employed	Mobility Analysis	No segmentation for low-educated, but segmentation for high-educated workers. Probability of formal job increases with wage differential.
Tannuri-Pianto and Pianto (2002)	Brazil	Non-registered workers	2-step Heckman selection Model	Positive selection for non-registered workers and negative selection for registered workers. Probability of formal job increases with wage differential.
Pratap and Quintin (2003)	Argentina	Non-registered workers	Propensity Score Matching	No Evidence of Segmentation.

Table 2.1: Summary of Studies on Formal x Informal Wage Differential and Segmentation

and Henley (2001) estimate an endogenous switching regression model for informal (non-registered) and formal (registered) workers and find evidence that informal sector workers allocate themselves into this sector due to comparative advantages, i.e., the more likely the individual is to choose the informal sector, the higher his/her expected wage in that sector is. However, they fail to find this comparative advantage result for formal sector workers. Tannuri-Pianto and Pianto (2002) use a quantile regression framework and find that comparative advantages - i.e. unobservables that lead to selection into the informal sector have a positive effect on earnings - play an important role in the case of informal sectors workers at the bottom of the distribution, whereas at the top of the distribution the effect is negative. Similarly to Carneiro and Henley (2001), they find that unobservables that make formal (registered) sector workers more likely to join the formal sector lead them to earn less than what would be expected based on their observed characteristics, i.e, they do not select themselves into formal jobs due to comparative advantages. They also show that the earnings gap favouring formal sector workers in the bottom quantile of the distribution is higher than at the top and cannot be explained only by differences in attributes. Differently, the gap at the top of the distribution is mostly due to differences in attributes[6].

One of the main criticisms of the cross-section studies on segmentation based on the analysis of the coefficient of the Inverse Mills ratio is that they usually rely on somewhat disputable exclusion restrictions in order to identify the sector allocation and wage equations separately[7]. Some authors have argued that the question of barriers to mobility into the formal sector appears to be an alternative way to settle the controversy on segmentation[8]. The studies surveyed by Behrman (1999) show that despite the relative high mobility between the two sectors, those who move from the informal sector to the formal sector have higher wage gains than the ones who move in the opposite direction. Gong et al. (2000) find that the probability of formal sector employment in Mexico increases with education level and that informal sector jobs are held by those with low family income, who cannot afford not to work at all. Differently, Maloney (1997) argues that the procyclical feature of the informal urban sector in Mexico makes it hard to affirm that it behaves as a "cushion" for bad times as stated by the traditional view of

[6]These results are in sharp contrast with the studies of Gong and Van Soest (2002) for Mexico and Pratap and Quintin (2003) for Argentina, that found no evidence of segmentation.

[7]Another criticisms refer to the assumption of the normality of the residuals in the participation equation and to the difficulty in measuring earnings in the informal sector, particularly, when it includes the self-employed and small employers.

[8]See Maloney (1997) and Maloney (1999).

segmentation in developing countries. Moreover, mobility patterns do not suggest a rigid labour market or one segmented along the formal/informal division (Maloney, 1999). Also using panel data for Mexico, Gong and Van Soest (2002) estimate a dynamic multinomial logit model with random effects for the choice of the sector, and two linear dynamic random effect equations for the wages in the two sectors. They find that the probability of formal sector employment strongly increases with wage differential. Their findings also suggest that for the lower educated workers, the dualistic view of the labour market is not a good description, since they would command a higher wage in the informal sector than in the formal sector. However, the most educated would command higher wages in the formal sector. In the case of Brazil, Barros et al. (1992) consider the mobility among three occupational states: registered (formal), non-registered (informal) and self-employed, and conclude that there is segmentation in the Brazilian labour market. Transitions to the registered sector always mean wage gains, whereas transitions from the registered sector to the non-registered sector or to self-employment always mean wage losses.

Pratap and Quintin (2003) estimations differ from the other studies because they do not correct wage equations for selectivity bias, on the contrary, they assume that all selection is controlled for using observables. They apply propensity score matching techniques in order to assess the wage differential between formal and informal workers in Argentina[9]. Their aim is to tackle the second source of bias in the wage differential studies as highlighted by Heckman and Hotz (1986): the assumption that the wage equation in correctly specified. They do not find any evidence of segmentation between formal and informal sectors in Argentina.

2.2 Formal and Informal Sector in Brazil: Some Stylized Facts

The literature on the informal sector in Brazil has basically three approaches. The first considers the informal sector as comprised of self-employed and small firm workers, the second considers informal sector as workers whose labour contract does not respect the labour code (non-registered), and the third puts together self-employed and non-registered workers. In this thesis we treat as informal sector only the second group. For this reason, we use the term registered and non-registered as synonymous with formal and informal sectors. In this section we will give an overview of what

[9]Informal workers are defined in their paper as workers who do not have access to some mandatory benefits.

happened to these two groups of workers during the 1980's[10] and 1990's, in order to set the scene for the other three chapters of this thesis.

2.2.1 Data

The data used in this chapter come from the Annual Household Survey (PNAD – *Pesquisa Nacional de Amostragem Domiciliar*) carried out by the Brazilian Statistics Office (IBGE – *Instituto Brasileiro de Geografia e Estatisticas*). We use data from 1981 to 1990, 1992, 1993 and from 1995 to 1999. There is no data for years when the national census is carried out, such as 1991, and in 1994 the survey was not conducted due to lack of funds. The representative sample consists of around 100.000 households covering the whole country with the exception of North rural area (Amazon area).

The main difficulty in working with the whole series of the PNAD is to filter the sample in order to disentangle non-registered workers from public (civil and military) servants for the period 1981 to 1988. As public servants do not have a registered work-card, they were classified as non-registered workers in the earlier surveys. Such problem did not happen in the surveys from 1989 onwards, because the individuals were directly asked whether they were public servants or not. In order to overcome this difficulty we filter possible public servants using the information on the worker's occupation and industry affiliation. To keep the consistency of the procedure, we ignored the actual information on the registration status available for the period 1989 to 1999 and applied the same filter we used for the 1981 to 1988 period.

2.2.2 Some Descriptive Statistics

One of the distinguishing features of the Brazilian labour market is the existence of a large number of workers whose job contract is not regulated by the legal labour code, i.e, they do not have a "signed work-card"[11]. These contracts are informal and illegal but make up something around 40% of the "wage workers"[12] and seem to have increased since the early 1990's.

Table 2.2 shows the mean and the standard deviation for several variables sep-

[10] For an excellent analysis of the changes in the informal sector - understood as self-employed plus non-registered workers - during the 1980's see Barros et al. (1993).

[11] The possession of a "signed work-card" (registration) gives workers several rights in terms of access to job-related public funds (e.g. unemployment benefit) and also to legally mandatory fringe benefits (e.g. paid vacations).

[12] Wage workers means remunerated employees.

arately for registered and non-registered workers[13] in three selected years: 1981, 1990 and 1999. Non-registered workers are more likely to be younger and less educated than registered workers, but these differences between the two groups have decreased over time, particularly, in relation to education. The average age of non-registered workers was 30(28) years in 1999(1981), whereas the registered average age was 33(31) years and the average years of schooling for non-registered workers was 5.9(3.1) and for registered, 7.8(6.2). Registered workers earned more than non-registered workers, but the gap narrowed considerably during the 1990's. The two groups used to work similar hours in 1981 and 1990, but in 1999 non-registered workers worked fewer hours than registered workers. However, the standard deviation of the working hours was much higher for non-registered workers. The presence of women increased in both groups, particularly, among the non-registered: in 1999 they represented 50% of the total non-registered workers. The regional distribution of the two groups remained quite stable over this period. Non-registered workers, however, seem to be over-represented in the Northeast.

The participation of the agricultural sector as a proportion of registered workers increased a lot between 1981 and 1999, while its participation in the non-registered sector (which was the largest in 1981) decreased[14]. Surprisingly, the manufacturing and the productive service sectors increased their participation in the pool of non-registered workers and reduced it in the pool of registered. The retail sector, the social services, and the lodging, food and other services expanded their participation among both registered and non-registered workers, whereas the constructing sector squeezed their participation among registered and remained relatively constant among non-registered. Despite being less prevalent in metropolitan areas, the participation of non-registered workers in those areas increased during this period. Differently, the participation of registered workers decreased in metropolitan areas. It seems that there were two opposite movements in this period: the rural sector, due to changes in agriculture, became less "informal", and the urban sector, due to the increase in informality in several sectors, became less "formal".

Data about race, unionization, tenure[15] and size of the firm were only available in more recent surveys. White workers are predominant among registered workers, whereas non-white workers are predominant among the non-registered. However, the proportion of whites among non-registered workers increased from 1990 to 1999. Non-registered workers were more likely to be employed by small firms. In 1990, only

[13]This sample excludes self-employed, non-remunerated workers, and public servants.

[14]The increase of the proportion of registered workers in the agriculture is due to the expansion of the highly productive agribusiness in the rural sector, replacing the traditional farming practices.

[15]Tenure is defined as time (in years and months) in the current job.

	1981		1990		1999	
	non-registered	registered	non-registered	registered	non-registered	registered
Years of schooling	3.07	6.17	3.82	6.87	5.85	7.78
	(3.15)	(4.12)	(3.40)	(4.16)	(3.63)	(3.86)
Gender (male=1)	0.65	0.72	0.63	0.68	0.49	0.61
	(0.48)	(0.45)	(0.48)	(0.47)	(0.50)	(0.49)
Age	28.01	31.06	28.28	31.86	30.26	32.87
	(13.96)	(11.06)	(13.83)	(11.06)	(12.68)	(10.78)
Experience (years)	18.94*	18.89	18.46	18.99	18.40	19.08
	(14.90)	(12.62)	(14.87)	(12.58)	(13.87)	(12.33)
Log hourly wage (R$ Sept. 1998)	-0.26	0.83	-0.36	0.60	0.04	0.67
	(0.81)	(0.85)	(0.88)	(0.91)	(0.81)	(0.76)
Hours	46.21	47.15	44.74	44.00	41.86	44.78
	(13.62)	(10.45)	(13.32)	(9.11)	(15.00)	(9.50)
Northeast	0.33	0.19	0.35	0.20	0.30	0.20
	(0.47)	(0.39)	(0.48)	(0.40)	(0.46)	(0.40)
North	0.06	0.06	0.08	0.07	0.08	0.04
	(0.23)	(0.24)	(0.28)	(0.26)	(0.28)	(0.20)
Southeast	0.35	0.45	0.30	0.44	0.34	0.43
	(0.48)	(0.50)	(0.46)	(0.50)	(0.47)	(0.50)
South	0.11	0.20	0.10	0.19	0.15	0.23
	(0.32)	(0.40)	(0.30)	(0.39)	(0.36)	(0.42)
Midwest	0.15	0.10	0.17	0.10	0.13	0.10
	(0.35)	(0.30)	(0.37)	(0.30)	(0.34)	(0.29)
Agriculture	0.35	0.03	0.27	0.05	0.19	0.06
	(0.48)	(0.18)	(0.45)	(0.22)	(0.39)	(0.23)
Manufacturing	0.09	0.35	0.12	0.33	0.11	0.27
	(0.29)	(0.48)	(0.32)	(0.47)	(0.31)	(0.44)
Constructing	0.11*	0.11	0.09	0.07	0.10	0.05
	(0.32)	(0.31)	(0.28)	(0.25)	(0.30)	(0.22)
Retail	0.08	0.14	0.10	0.16	0.11	0.17
	(0.27)	(0.35)	(0.30)	(0.37)	(0.31)	(0.37)
Lodging, food and other services	0.29	0.12	0.31	0.13	0.37	0.20
	(0.45)	(0.32)	(0.46)	(0.33)	(0.48)	(0.40)
Productive services	0.05	0.19	0.06	0.19	0.08	0.16
	(0.22)	(0.39)	(0.24)	(0.39)	(0.27)	(0.37)
Social Services	0.03	0.07	0.05	0.08	0.04	0.10
	(0.17)	(0.25)	(0.21)	(0.27)	(0.21)	(0.30)
Metropolitan area	0.31	0.63	0.30	0.55	0.43	0.55
	(0.46)	(0.48)	(0.46)	(0.50)	(0.50)	(0.50)
Race (white=1)			0.40	0.57	0.45	0.59
			(0.49)	(0.50)	(0.50)	(0.49)
Size (more than 10 = 1)			0.25	0.81	0.21	0.69
			(0.43)	(0.39)	(0.41)	(0.46)
Tenure (years)			2.96	4.79	3.12	4.78
			(5.98)	(6.02)	(4.95)	(5.63)
Union					0.02	0.24
					(0.12)	(0.43)
% earning less than mw	0.61	0.07	0.48	0.05	0.35	0.01
% earning the mw	0.01	0.04	0.06	0.09	0.13	0.08
% earning more than mw	0.39	0.90	0.47	0.86	0.53	0.91
N	38895	61840	27506	41261	25885	37706

(*) Mean of the variable is not statistically different between the two samples at 5% of significance

Table 2.2: Descriptive Statistics for Non-Registered and Registered Workers in Selected Years (1981, 1990 and 1999)

25% of non-registered workers worked in firms with more than 10 employees, this figure was down to 21% in 1999. This contrasts with the sample of registered workers: 81% of which worked in firms with more than 10 employees in 1990, however, in 1999 only 69% were in this situation. Non-registered workers had less seniority than registered workers and were extremely less likely to be unionized: only 2% of non-registered workers were unionized, whereas 24% of registered workers were unionized in 1999.

Overall non-registered workers have a significant disadvantage in terms of productive attributes. They are less educated, more likely to be employed in smaller firms and in low productivity sectors. They are also less likely to be unionized and more prone to be discriminated against since female and non-white workers are over-represented among them.

On the bottom of Table 2.2 we report the proportion of workers whose wage was lower, equal and higher than the minimum wage. Non-registered workers exhibit a high degree of non-compliance, however, non-compliance decreased considerably during the period under investigation[16]. At the same time, there was a rise in the proportion of non-registered workers whose wages were equal to the minimum wage: 13% in 1999 compared to 1% in 1981[17].

Overall it seems that changes in the composition of the two groups between 1981 and 1990 may explain at least part of the fall in the wage differential between the two groups. In the next subsections we will analyze the effects of these changes more thoroughly.

2.2.3 Changes in Employment Structure during the 1980's and the 1990's

In the 1980's the proportion of registered workers in the occupied population followed very closely the behaviour of the business cycle. During the recession of the early 1980's it decreased sharply, but after 1984 as the economy recovered, it slightly increased. The recession of the early 1990's led to another reduction in the proportion of registered workers, but this time, even with the recovery of the economy after 1993, the proportion of registered workers did not react. In contrast, the propor-

[16]Notice that in 1999 the non-compliance among registered (non-registered) workers was down to 1% (35%), compared to 7% (61%) in 1981.

[17]Employers caught not paying the minimum wage or other labour regulations are in general fined and sued, however, this is not a common practice in the country. The high figures on non-compliance with the minimum wage legislation in the 1980's cannot be attributed to measurement error. If any error like that was to be expected, it would be on reporting minimum wage earnings since it has been used as an index for a long time.

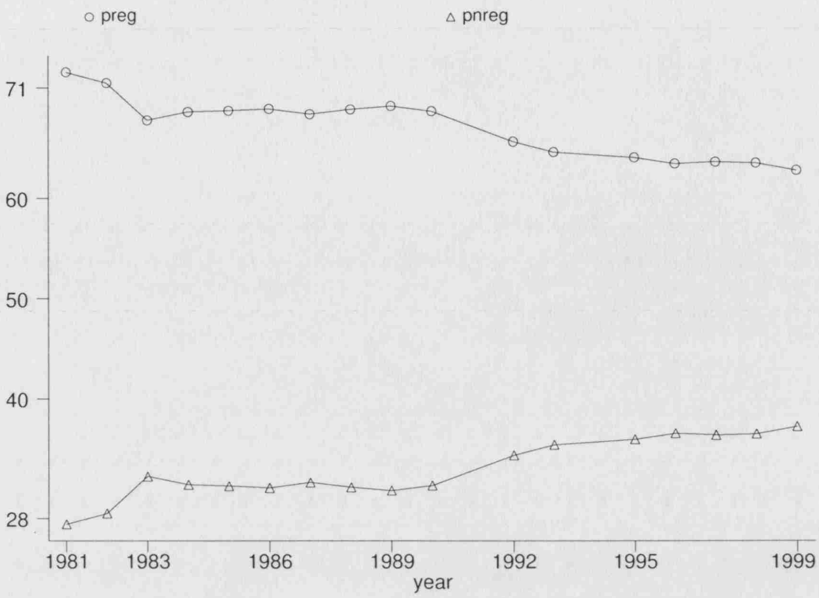

Figure 2.1: Proportion of Registered and Non-Registered Workers (in %) - 1981 - 1999

tion of non-registered had a counter-cyclical behaviour - as expected by the buffer interpretation of the informal sector - during the early 1980's, peaking in 1983 and reaching its lowest level in 1990. After that, the proportion of non-registered workers increased slightly and has remained rather constant since 1995[18].

The increase of the proportion of non-registered workers was more intense in the non-agricultural industries. In fact, the agriculture industry experienced an increase in the proportion of registered workers, in spite of its initial and still high level of non-compliance (see Table 2.2). However, since the proportion of agricultural sector jobs has decreased continuously over time, the aggregate figure is dominated by the changes in the non-agricultural sector. Looking only at this latter group the proportion of non-registered workers increased from 30% in 1981 to 40% in 1999 (see Figure 2.1), and the bulk[19] of this increase was concentrated after 1990, just after the country started the market-oriented reforms, such as the programmes of privatization and the process of trade liberalisation.

While the proportion of non-registered workers increased, the raw wage gap between registered and non-registered workers fell between 1981 and 1999. As shown

[18]The increase in the number of self-employed seems to have accounted for the major part of the decrease in the proportion of registered workers. However, as the reduction over the period was higher for the proportion of registered workers than for non-registered workers, this led to a lower proportion of registered workers in the pool of employees.

[19]We are not taking into consideration here the isolated peak observed in 1983 due to the severe recession observed then.

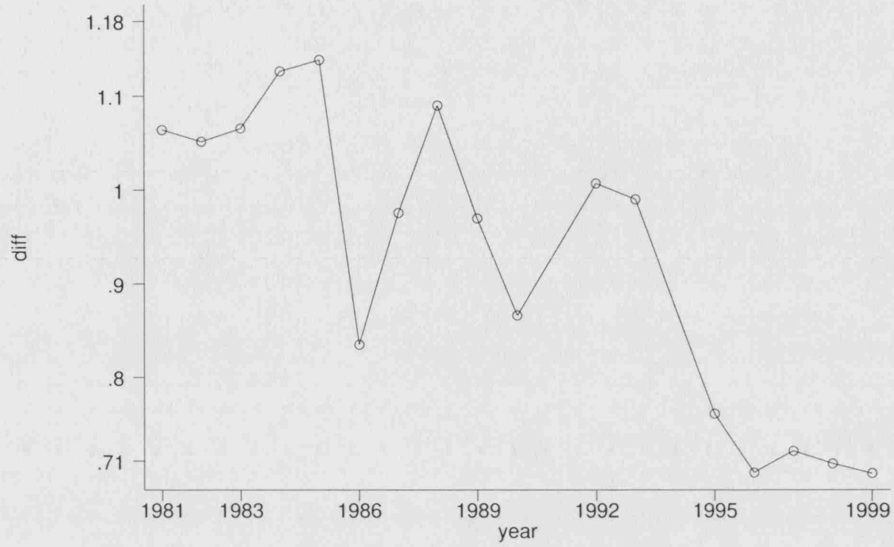

Figure 2.2: Wage Premium for Registered Workers - 1981 - 1999

in Figure 2.2, in 1981 the raw ratio of log real hourly-wage between registered and non-registered workers was 1.08, but in 1999 it was down to 0.71 [20]. Many factors may have triggered such decrease: composition effects (due to the improvement of non-registered workers productive attributes in comparison to their registered counterpart), higher returns to attributes in the non-registered sector due to changes in the economic environment, and so on. In order to have a clear view of what happened with the wage gap once one controls for the observable characteristics of workers and firms, the following subsections will discuss the main results of a set of regressions for registered and non-registered workers and present some decomposition exercises for changes in average wages and changes in wage inequality.

[20] It is true that there is a major dip in 1986 and 1990, but this episodical movements can be explained by the effect of two unorthodox plans: *Plano Cruzado* in 1986 and *Plano Collor* in 1990. The *Plano Cruzado* froze wages and prices in February 1986. As wages and prices in the informal sector are not easily controlled as wages and prices in the formal sector, the increase in the demand for non-tradables observed in that period benefited informal workers both in terms of employment and wages, leading to a sharp fall in the wage differential between formal and informal workers, see Camargo and Ramos (1988) for a discussion of this point. As for the Plano Collor, the dip was due to the recession that followed the freezing of assets that aimed to control inflation. The recession curbed wage rises in the formal sector reducing the wage gap.

2.2.4 Evolution of Wage Differential Between Registered and Non-registered Workers

All regressions in this subsection are run for workers who worked at least 20 hours[21] in the week immediately before the interview, had positive earnings, were not employed in farming activities, and were between 14 and 65 years[22]. In order to check different patterns of segmentation according to gender we run separate regressions for men and women. In the same vein, wage equations for registered and non-registered workers are run separately so that we can follow the evolution of the gender wage gap and of the returns to skills for both groups.

The dependent variable in all specifications is the log of hourly real wage[23] and the regressors are group of years of schooling (illiterate – $yos1$, some primary – $yos2$, complete primary and some elementary – $yos3$, complete elementary and some secondary – $yos4$, complete secondary and some college – $yos5$, and complete college and post-graduation – $yos6$)[24]; potential experience $(age - years of schooling - 6)$[25]; potential experience squared; dummies for four regions; dummy for metropolitan area and, when appropriated, dummies for gender; work-card (registration); size of the firm, tenure and race.

To assess the consistency and accuracy of filter used to separate public servants from non-registered workers we re-estimate the wage equations for the 1989-1999 period using the available information on the actual public servant status. The parameter estimates are quite close to the ones yielded by the filtered data, but the sample size turns out to be larger. This means that the filter underestimates the number of both non-registered and registered workers. It classifies some registered

[21]In 1982 workers working less than 20 hours corresponded to 1.5% of the full sample (registered and non-registered), in 1999 it was up to 3.7%. Workers younger than 16 years old were 4.58% among employees working less than 20 hours and 3.37% among those working more than 20 hours, the figures for 1999 were 2.43% and 1.94%, respectively.

[22]The legislation on the minimum age for which is allowed to work has changed over the sampled period. We restrict the analysis to the age bracket 14 to 65 years old because 14 is the legal minimum wage to work as an apprentice. This low age is not a worry for the results as the filter of 20 hours will prevent youngest workers without a strong attachment to the labour market, as measured by weekly hours of work, of being selected into the sample.

[23]The nominal wages were deflated by the INPC (Consumer Price Index) based on September 1998.

[24]We choose to enter the education variable as groups and not as a continuous variables in order to capture the non-linearities of the return to education and to have a clear picture of the evolution of the return for different groups.

[25]Years of schooling is measured as the number of years that corresponds to the grade achieved by the worker, it does not take into account repetition or age/grade delays. Therefore, despite measuring education in a proper way, it can carry some measurement error to potential experience as years of schooling minus 6 does not gauge accurately the age at which the worker entered the labour market.

or non-registered workers as public servants, which yields a smaller sample of private sector employees. Another difference between the two samples is that for the 1989 to 1999 sample, it is possible to use a broader set of regressors, such as tenure, size and race variables[26]. It is worth mentioning that the "size of the firm" variable, which unfortunately has to be coded as a binary (more than ten = 1), has an important effect on the estimation. Once it is included in the regression, the magnitude of the coefficient for the dummy for a registered work-card is drastically reduced[27].

Figure 2.3 shows that the log hourly real wage for both registered and non-registered workers followed a similar path over time. The differences are concentrated in the period 1988-1993, when the average wage for the registered workers declined moderately, whereas the average wage for the non-registered increased from 1988 to 1990 and then decreased continuously until 1993. Despite sharing the same trend, the intensity with which each group's wage react to changes in the business cycle varied a lot. The wage recovery after 1993, for instance, was sharper for the non-registered than for the registered workers[28]. This pattern led the ratio of log real wage between registered and non-registered workers to decline from 1992 onwards as shown in Figure 2.2.

According to the regressions based on the filter-based sample, the 1990's witnessed a lower volatility of wage differential[29] when compared to the 1980's, in particular, the differential seems to have stabilised around 36% since 1995[30] (Figure 2.4). This threshold contrasts sharply with the peaks observed in 1985, 1988 and 1992, when the wage differential reached something around 70%, after controlling for age, education, region, metropolitan area, potential experience and gender. However, the evolution of the controlled wage differential is quite similar to the one observed for the raw wage differential, in particular, the dips observed in 1986, 1990 and 1995 are not attenuated.

The second classification, which permits to control for self-declared public servants instead of using our occupation/industry filter, yields similar results to the ones presented above (see Figure 2.5). The wage differential is somewhat smaller

[26]Union information in only available for 1988 and from 1992 onwards.

[27]The sample size of the specification that controls for size of the firm is lower than the others because we drop all "domestic servants" of the sample. This procedure is justified because almost all domestic servants would assume the value 0 (less than 10) for the variable firm size, which could bias the results.

[28]The peak for both registered and non-registered workers observed in 1986 was due to the *Plano Cruzado*, an unorthodox stabilization plan, which froze wages and prices.

[29]Wage differential is measured as the antilog of the coefficient of a dummy variable coded 1 for registered workers and 0 for non-registered workers.

[30]This figure is half of raw wage differential that was around 70% since 1996 as shown in figure 2.

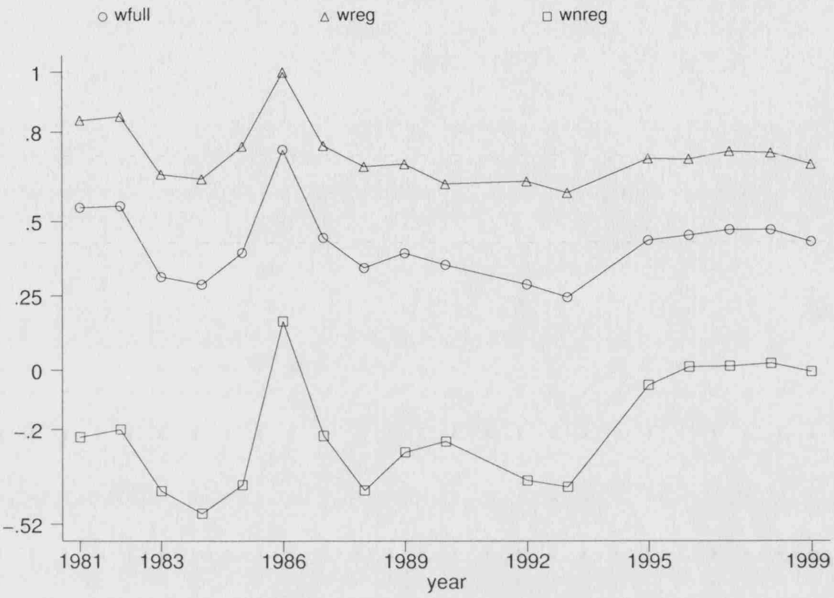

Figure 2.3: Log Hourly Real Wage Rate - 1981 - 1999

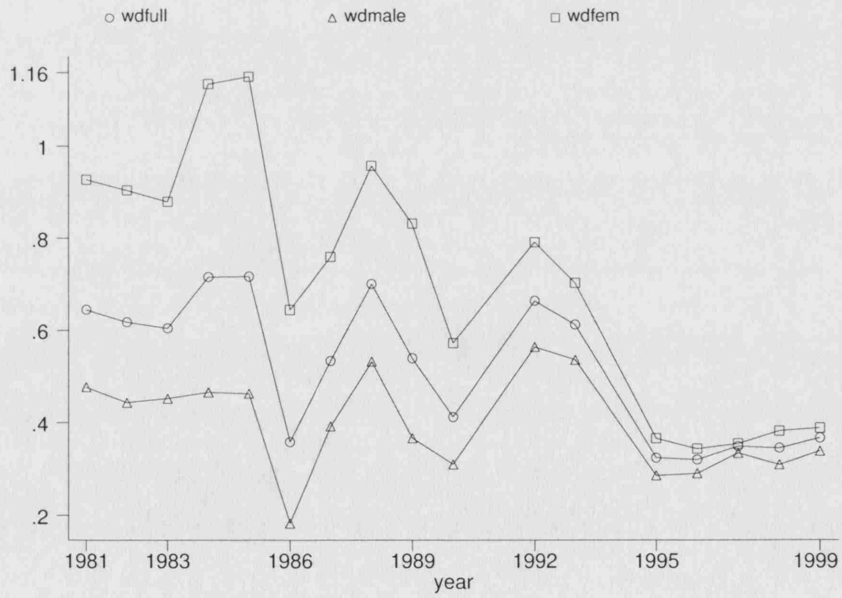

Figure 2.4: Wage Differential between Registered X Non-registered (Filter based) - 1981 - 1999

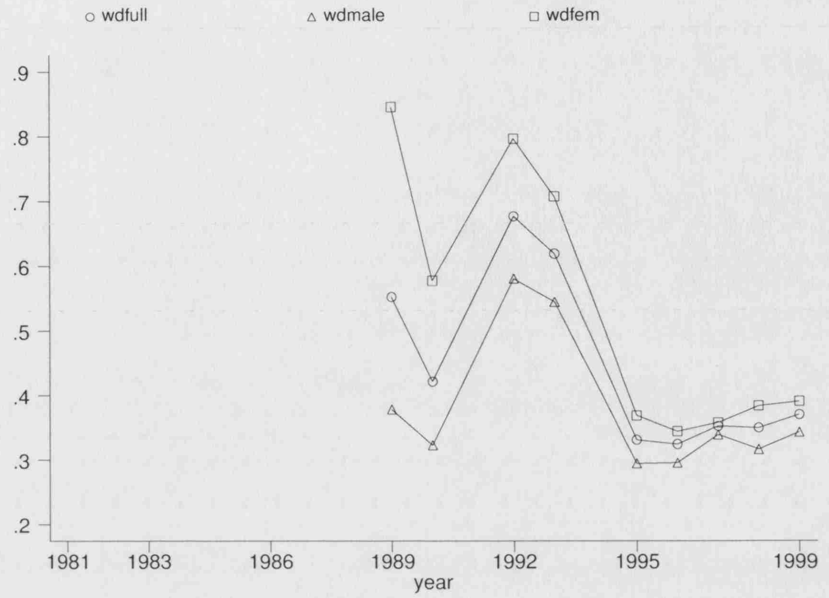

Figure 2.5: Wage Differential between Registered X Non-registered (Survey based) - 1989 - 1999

for more recent years in the survey-based classification, but the differences are very small to shed doubts on our filter-based classification.

The third set of results uses the same classification of the second, but exploits more controls that were made available only in the more recent surveys. As expected, the inclusion of size of the firm, tenure and race has a huge impact in the controlled wage differential. Figure 2.6 shows that the estimated wage differential is almost halved when one includes those variables. However, its pattern over time is exactly the same as shown on the previous results: a fall in 1990 and 1995 and a peak in 1992. According to this specification the wage differential would be around 18% in 1999.

As noticed in the previous subsection, the increase in the participation of women in both registered and non-registered sectors was one of the main changes observed during this period. Figures 2.7 to 2.9 show that the wage differential between male and female workers decreased substantially over time for both registered and non-registered workers. However, whereas the non-registered sector seems to have a higher wage gap in favour of male workers for the specifications with fewer controls (Figure 2.7 and 2.8), in the specification with more controls (Figure 2.9) the registered worker sample is the one that seems to have a higher wage gap in favour of male workers.

Two striking features are revealed by the two sets of figures considered above.

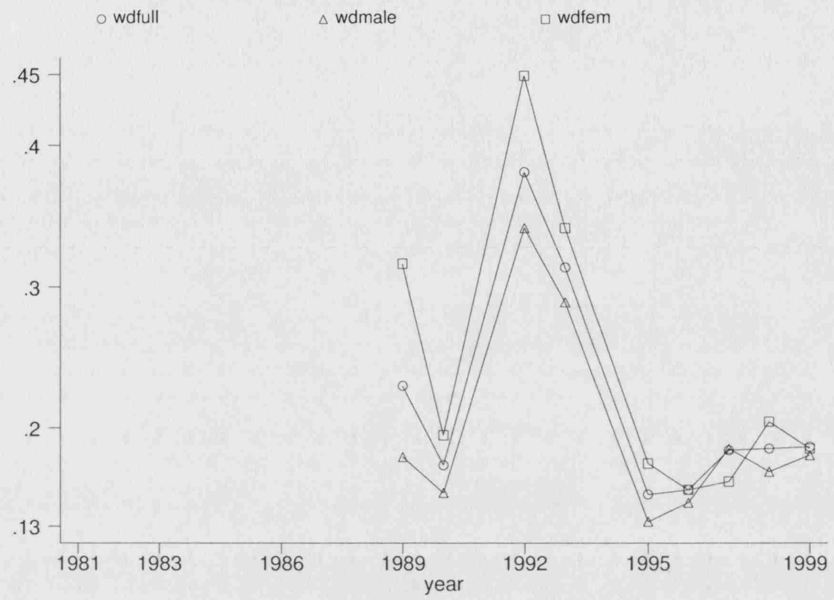

Figure 2.6: Wage Differential between Registered X Non-registered (Survey based and more controls) - 1989 - 1999

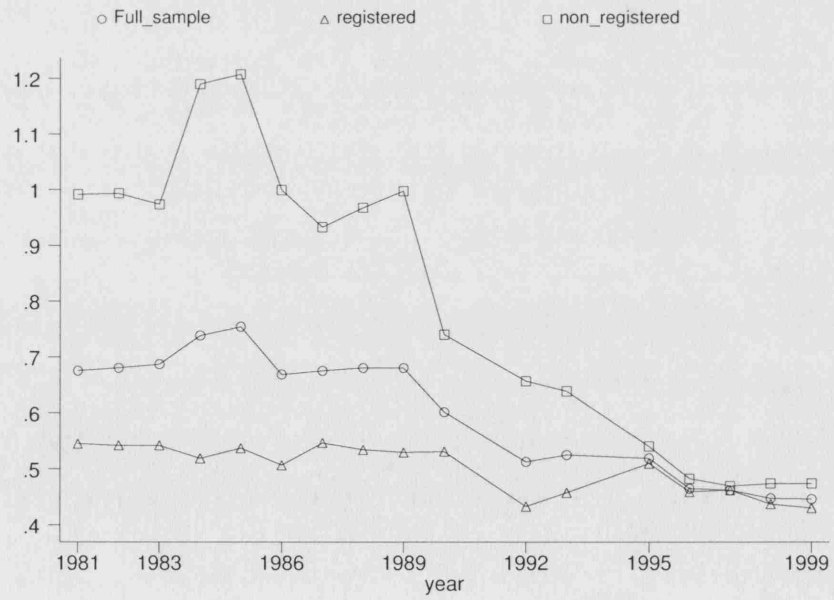

Figure 2.7: Wage differential between Male X Female (Filter based) - 1981 - 1999

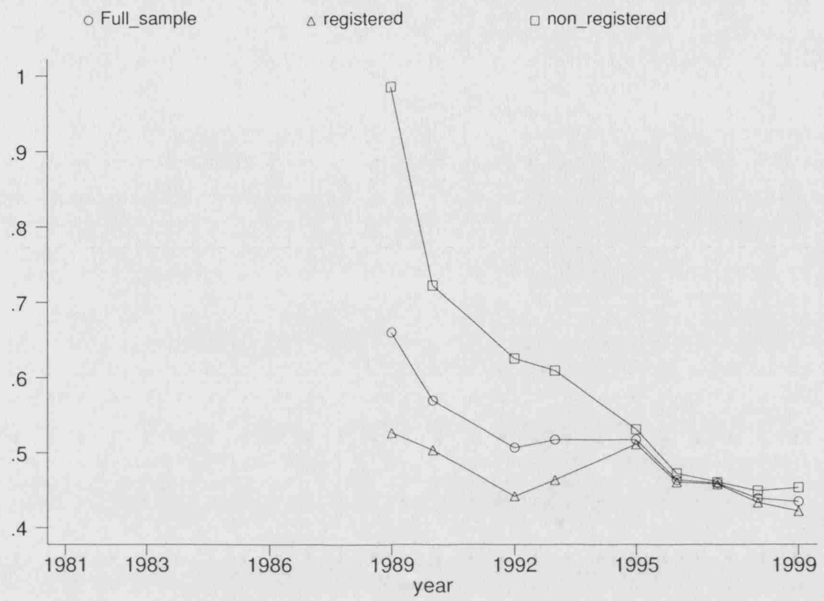

Figure 2.8: Wage differential between Male X Female (Survey based) - 1989 - 1999

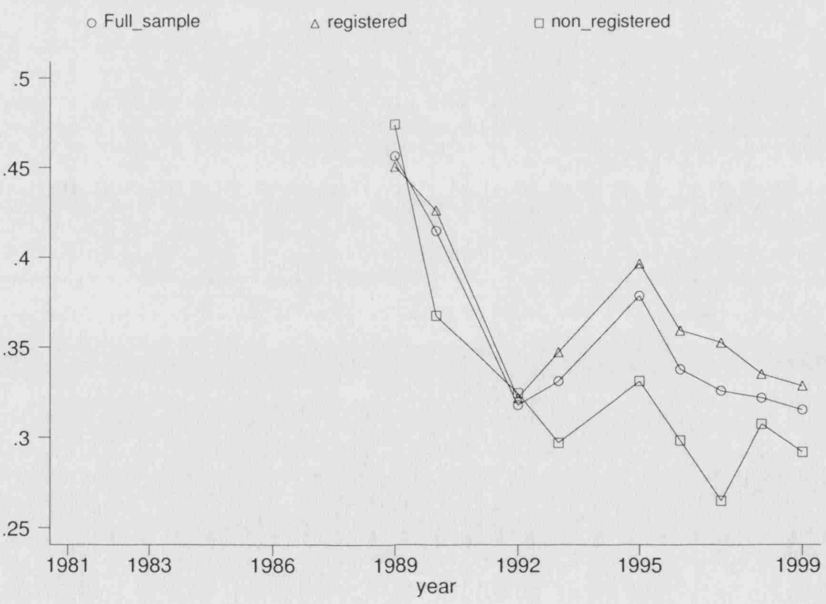

Figure 2.9: Wage differential between Male X Female (Survey based and more controls) - 1989 - 1999

First, the wage differential among female workers was considerably higher than among male workers. However, such discrepancy has diminished over time and at the end of the sample period both measures were quite similar (Figures 2.4 to 2.6). Second, the gender wage gap was higher for non-registered workers than for registered workers. But here again, there has been some convergence over time (Figures 2.7 to 2.9).

Looking at the wage premium over the wage distribution, a simple quantile regression for key years 1982 and 1999 reveals that the wage premium has fallen for the 10th, 25th, 50th, 75th, and 90th percentile with almost the same pace for all groups. However, it is quite remarkable that the wage premium is much higher for the lowest percentiles. It seems that the wage premium is much more important for workers at the bottom of the distribution and this feature has not changed over time[31]) in 1999.

2.2.5 Returns to Education for Registered and Non-Registered Workers

The wage premium[32] for workers with complete college was quite stable during the early and middle 1980's. It peaked in 1988 and then started falling until 1992 to a level lower than the 1980's average. However, after 1992, the wage premium for college workers started increasing and achieved a level above the peak observed in 1988[33] (Figure 2.10). Green et al. (2001) argue that such evidence is in line with the hypothesis that, somehow, the trade reform in the early 1990's has triggered an increase in the returns to education for high-skilled workers[34].

The complete high school or some college group was the big loser in this period. Its wage premium, over and above the complete elementary and some high school group (yos5), witnessed the sharpest decrease among all groups. Surprisingly, the illiterate group gained some ground and managed to reduce its wage differential in comparison with the group with incomplete primary education. The relative good performance of the lower education group is also found by Green et al. (2000) for all

[31]The wage premium for the 10th (90th) percentile was 95% (48%) in 1982 was down to 50% (14%

[32]These figures depict the wage premium for each education level over and above the group immediately below. For that reason, the reference group for the group some elementary education (yos2),i.e, the illiterate group (yos1) does not appear in the graphs.

[33]The conditional wage differential between college workers and complete high school or some college was 176% in 1999. In 1992, this differential was around 132%, the lowest level in the sample.

[34]There is still no strong evidence of the direct links between trade liberalisation and increase in returns to education in Brazil. We will discuss this hypothesis in the chapter 4 of this thesis.

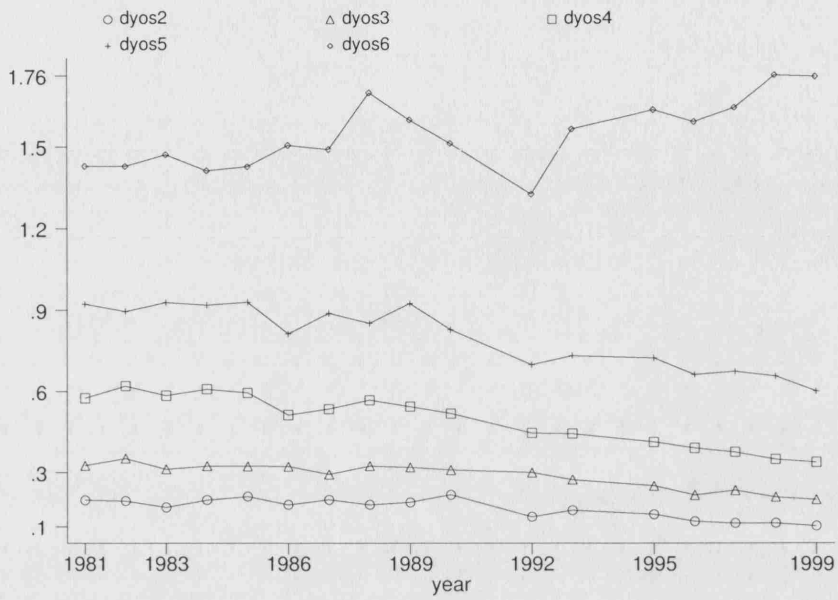

Figure 2.10: Relative Returns to Education (Full Sample) - 1981 - 1999

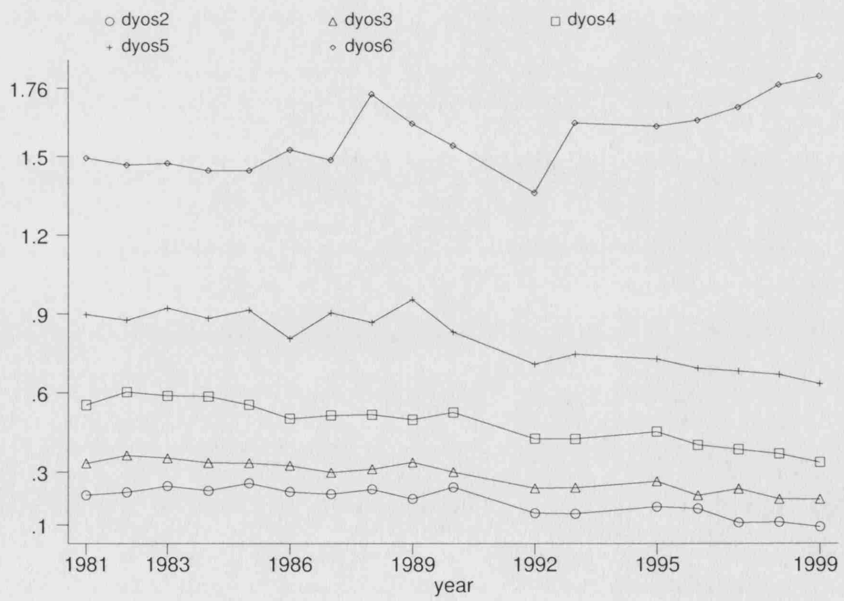

Figure 2.11: Relative Returns to Education (Registered Sample) - 1981 - 1999

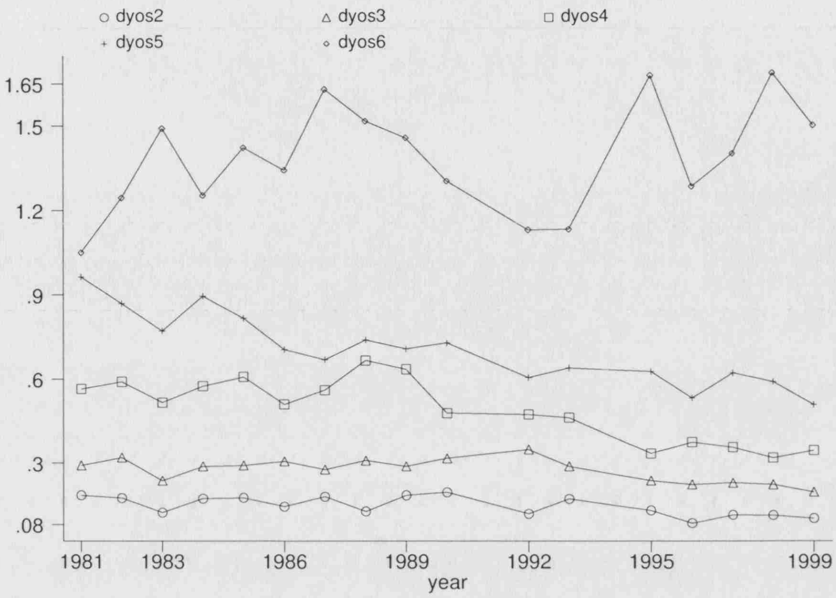

Figure 2.12: Relative Returns to Education (Non-Registered Sample) - 1981 - 1999

occupied population (including self-employed and agricultural sector) for Brazil[35] and by Behrman et al. (2001) for a panel of 18 Latin American countries. These two papers also report the increase in the premium for college education relative to secondary education. Fernandes and Menezes-Filho (2000) using Brazilian data conclude that the fall in the relative returns to education for all groups - with the exception of the college group - was the main factor triggering the reduction in wage inequality between 1983 and 1997[36].

Looking now at the returns to education for registered and non-registered groups separately, one can see that the aggregate pattern is determined by the behaviour of the returns to education for registered workers (Figure 2.11). The wage premium for non-registered workers (Figure 2.12) with some college is somewhat more volatile than the one observed for their registered workers counterpart. Its premium over and above the group with secondary or incomplete college is lower, having oscillated within the range of 100% and 150%. Nevertheless, despite the lack of a continuous pattern, it seems that there has been an increase in their wage premium after 1993,

[35]The authors attribute this result to the reduction in the supply of illiterate workers over time in Brazil.

[36]It is important to notice however, that the returns to education vary a lot over the wage distribution. Arabsheibani et al. (2003) show that the returns to education are much higher in the top of the distribution. Similar results were also found when we run quantile regression on this sample. Interesting the college group was the only group that improved over the group immediately below between 1992 and 1999, but only from the median up the distribution.

one year after the increase observed for registered workers group[37].

The same picture is found when one looks at the sample from 1989 to 1999 using the self-declared public servant status filter to separate out public servants from the pool of non-registered workers. Similarly, the third specification – controlling for size of the firm, race and tenure – displays lower returns to education but does not show any relevant difference relative to the pattern observed in the former specifications[38].

2.2.6 Decomposing Wage Differential and Accounting Inequality for Registered and Non-registered Workers: 1981-1999

In this subsection we will evaluate how changes in the average wage for the full sample and also for registered and non-registered workers can be decomposed into changes in attributes and in returns to attributes through Oaxaca-Blinder decomposition. We will also decompose the wage gap between registered and non-registered workers into differences in attributes and in their returns. Additionally, we will use Juhn-Murphy-Pierce (1991) methodology to investigate the determinants of the narrowing in the wage gap.

As the fall in the wage gap between registered and non-registered workers may have triggered a decrease in the overall wage inequality, we will also investigate how changes into the wage premium for registered workers and in its size have affected inequality. We will apply Fields' (2002) method to decompose both the level of income inequality and how it changed over time for the full sample and for registered and non-registered workers separately. The measure of inequality used here is the log variance of the wages. Additionally, we will apply Juhn, Murphy and Pierce (1993) decomposition in order to look at inequality at different parts of the wage distribution. This is important because non-registered workers are over-represented on the lower tail of the wage distribution.

[37]Data not shown here reveal that for non-registered female workers with complete college the increase in the relative return to education for college workers after 1993 does not compensate the fall after 1987. The high returns observed in the early 1980's for this group may be due to its very small sample size. This appears to be the only major difference regarding gender pattern in returns to education.

[38]See Figures A.1 to A.6 in the Appendix.

Oaxaca-Blinder Decomposition and the Juhn-Murphy-Pierce (1991) Extension

A simple Oaxaca-Blinder decomposition may shed some light in order to understand the different patterns in the evolution of the average wage for the full sample and for the registered and non-registered workers separately. The ratio of log hourly real wage decreased –0.092 log points between 1999 and 1981. A somewhat sharper decrease was observed for registered workers, -0.145 log points, whereas non-registered workers experienced an increase of 0.263 log points. Throughout the sample period the average wage of the non-registered workers increased more than the average wage of the registered in expansionary periods, such as 1981-1986 and 1992-1995, and decreased by a lower amount during recessions, such as 1987-1990. This different pattern led to a lower degree of segmentation as measured by the wage gap between registered and non-registered workers.

This pattern raises the question of whether non-registered workers have been improving their relative position due to changes in their observable attributes – a composition effect, probably caused by workers displaced from registered jobs getting a non-registered one – or due to an increase in the returns of their observable attributes. A second point is which of these components was most important in driving the fall in the wage gap between registered and non-registered between 1981 and 1999.

As for changes over time, the Blinder-Oaxaca decomposition[39], Table 2.3[40] reveals that changes in the composition of the characteristics helped to improve the average earnings, whereas returns to these characteristics have exerted an opposite force for all three samples: pooled, registered and non-registered[41].

As for the determinants of the average wage differential between the two groups, Table 2.4 shows that regardless of the weight scheme[42] used in the calculation, the

[39]The specification for the full sample includes dummy for registered workers. The other regressors are the same as the ones used in our basic specification based on the occupation/industry filter.

[40]In Table 2.3, panel A brings the results using the mean of the attributes of 1999 as weight for the change in coefficients, and the returns of 1981 as the weight for changes in the mean of the attributes, whereas panel B brings the results using the mean of attributes of 1982 as weight for the change in coefficients, and the returns of 1999 as weight for changes in the mean of the attributes.

[41]Actually only for non-registered workers the joint effect of changes in the constant and changes in return had a positive effect on average earnings, but all the positive effect came from a larger constant.

[42]In panel A, the means of attributes for the registered sample were used as weight for the difference in the coefficients and the coefficients of non-registered worker equation were used as weight for the difference in attributes. In panel B we reverse the weights.

		(A)			(B)		
	Δlwh	X_{99}*(β_{99}-β_{81})		β_{81}*(X_{99}-X_{81})	X_{81}*(β_{99}-β_{81})		β_{99}*(X_{99}-X_{81})
		Δconstant	Δreturns	Δmeans	Δconstant	Δreturns	Δmeans
Full Sample	-0.091	0.334	-0.513	0.088	0.3342	-0.5545	0.1290
		-366%	563%		-366%	608%	
			196%	-96%		241%	-141%
Registered	-0.145	0.0542	-0.3324	0.1328	0.0542	-0.3896	0.1900
		-37%	229%		-37%	268%	
			191%	-91%		231%	-131%
Non-registered	0.263	0.5193	-0.4622	0.2061	0.5193	-0.5634	0.3072
		197%	-176%		197%	-214%	
			22%	78%		-17%	117%

Table 2.3: Oaxaca-Blinder Average Wage Decomposition: 81-99

		(A)			(B)		
Segmentation	Δlwh	X_r*(β_r-β_{nr})		β_{nr}*(X_r-X_{nr})	X_{nr}*(β_r-β_{nr})		β_r*(X_r-X_{nr})
		Δconstant	Δreturns	Δmeans	Δconstant	Δreturns	Δmeans
1981	1.115	0.896	-0.352	0.570	0.896	-0.445	0.664
		80%	-32%		157%	-78%	
			49%	51%		40%	60%
1999	0.706	0.431	-0.108	0.384	0.431	-0.116	0.392
		61%	-15%		112%	-30%	
			46%	54%		45%	55%

Table 2.4: Oaxaca-Blinder Wage Gap Decomposition: 81-99

difference in attributes is more important in explaining the wage differential between registered and non-registered workers than differences in returns[43]. However, the different weight schemes give different results for the importance of each component on changes between 1981 and 1999, whereas in 1981 according to the weights used in panel A (panel B), 49% (40%) of the average wage differential between registered and non-registered workers was "explained" by differences in returns, in 1999 this proportion had decreased (increased) slightly to 46% (45%).

Juhn, Murphy and Pierce (1991) expand the simple Oaxaca-Blinder decomposition in order to take into account changes in the residual distribution. Their approach allows one to decompose changes in the wage gap between the formal and informal sector into changes in the observable components and changes in the unobservable components. The wage equation for formal and informal workers can be written, respectively, as:

$$W_{fit} = X_{fit}\beta_{ft} + \sigma_{ft}\theta_{fit} \qquad (2.1)$$

$$W_{iit} = X_{iit}\beta_{it} + \sigma_{it}\theta_{iit} \qquad (2.2)$$

where σ_{ft} and σ_{it} are the within-group standard deviation of wages in the formal and informal sectors in year t and θ_{fit} and θ_{iit} are the standardised residuals of each wage equation: θ_{fit}=$\epsilon_{fit}/\sigma_{ft}$. The wage gap between formal and informal sector workers is:

[43]Verry and Araujo (1996) found similar results.

$$D_t = \bar{W}_{ft} - \bar{W}_{it} = (\bar{X}_{ft} - \bar{X}_{it})\hat{\beta}_{ft} + \hat{\sigma}_{ft}\bar{\Delta\theta}_t \tag{2.3}$$

where $\bar{\Delta\theta}_t$ is the mean difference in the average standardized residual for workers in the formal and informal sector. Then, changes over time in the wage gap can be decomposed as:

$$\begin{aligned} D_t - D_{t-1} = [(\bar{X}_{ft} - \bar{X}_{ft-1}) - (\bar{X}_{it} - \bar{X}_{it-1})]\hat{\beta}_{ft} \\ +(\bar{X}_{ft-1} - \bar{X}_{it-1})(\hat{\beta}_{ft} - \hat{\beta}_{ft-1}) \\ +\hat{\sigma}_{ft}[\bar{\Delta\theta}_t - \bar{\Delta\theta}_{t-1}] + \bar{\Delta\theta}_{t-1}(\hat{\sigma}_{ft} - \hat{\sigma}_{ft-1}) \end{aligned} \tag{2.4}$$

The first term captures the effect of changes in the quantity of the observables, $X's$, the second term captures the effect of changes in the prices of the observables. The third term is called the "gap effect" and measures the effect of the changes in the relative position of informal workers in the formal wage distribution, i.e, it captures what would happen if the residual formal sector wage inequality were held constant between $t-1$ and t, but the percentile ranking of the informal wage residual had changed. If informal workers had moved up this distribution it can mean that they had increased their stock of unobserved characteristics or that they are less "discriminated" against. However, as being an informal sector worker is not the same as being "black" or "woman" in the labour market, since they do not have this "permanent" and "immutable" characteristics, it is hard to talk about a lessening in discrimination. It is much more likely that there had been some change in demand that somehow makes their "unobservable" characteristics more valued in the labour market. The last term is the so-called "unobserved prices" effect and measures the change in the wage gap due to the changes in inequality among formal sector workers. It means that a rise in inequality (over time) would increase the wage gap between formal and informal sector workers, even if the percentile ranking of informal sector wage residual had not changed over this period.

The results of the JMP decompostion in Table 2.5 show that all components contributed to the narrowing of the wage gap. However, most of the reduction in the wage gap was due to the "gap effect", 52%. According to our interpretation this is a sign that returns to observables and the improvement on the productive endowments of informal workers were much less important than changes in the economic environment either via demand shocks or supply shocks that are not readily

D_{99}	0.697
D_{81}	1.064
$D_{99} - D_{81}$	-0.367
Observable	
Quantity	-0.080
Prices	-0.081
Unobservables	
Gap	-0.191
Unobservable Price	-0.015

Table 2.5: JMP Decomposition of Changes in the Wage Gap between Formal and Informal Workers: 1981-1999

observable. It is not clear what sort of unobservables could have triggered this result. Among the hypotheses that will be assessed in the next chapters are the impact of the trade liberalisation and the minimum wage indexation of the informal sector wages.

Accounting for Inequality and Decomposing its Changes

Figure 2.13 shows that the standard deviation for the log hourly real wage of registered workers used to be lower than the standard deviation for the non-registered workers (with the exception of 1983), but from 1995 onwards, the latter has fallen sharply and has remained lower than the standard deviation for registered workers. Given the fall in the wage gap between registered and non-registered workers and the decrease in the standard deviation of the non-registered worker's earnings (and also in a smaller scale for the registered workers), it is not surprising that the dispersion of log hourly real wage has diminished for the full sample.

The standard deviation of the residuals of the wage regressions for both registered and non-registered workers increased during the 1980's and decreased during the 1990's (Figure 2.14)[44]. The fall was sharper from 1995 onwards when the hyper-inflationary process was controlled. Nevertheless, it is striking that the "unexplained" dispersion of the log wages is higher for non-registered than for registered workers over the entire period.

To assess the importance of the falling wage differential through the 1980's and the 1990's for the overall reduction in the log variance of wages, we perform the

[44]Actually the graphs for the standard deviation of the log hourly real wage and for the standard deviation of the residual of the wage regressions are quite similar.

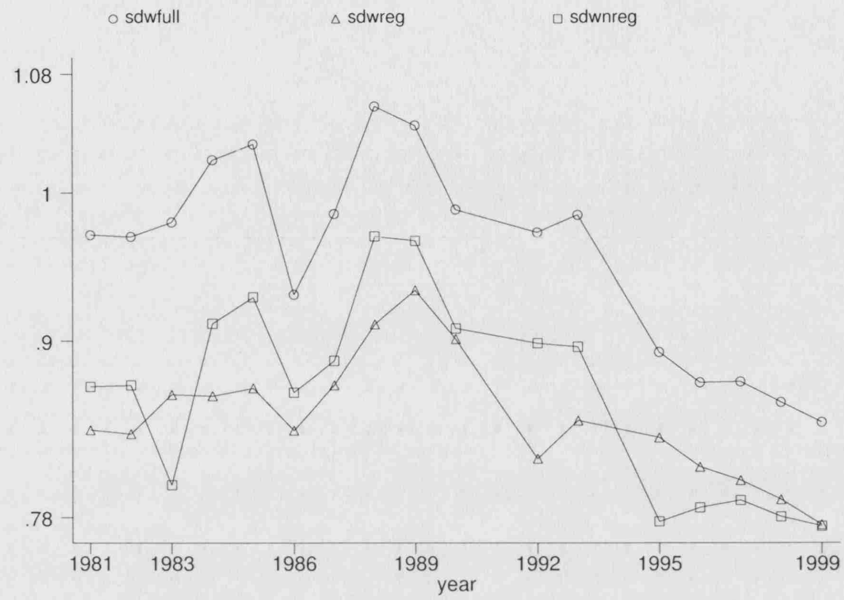

Figure 2.13: Standard Deviation of the Log Hourly Real Wage - 1981 - 1999

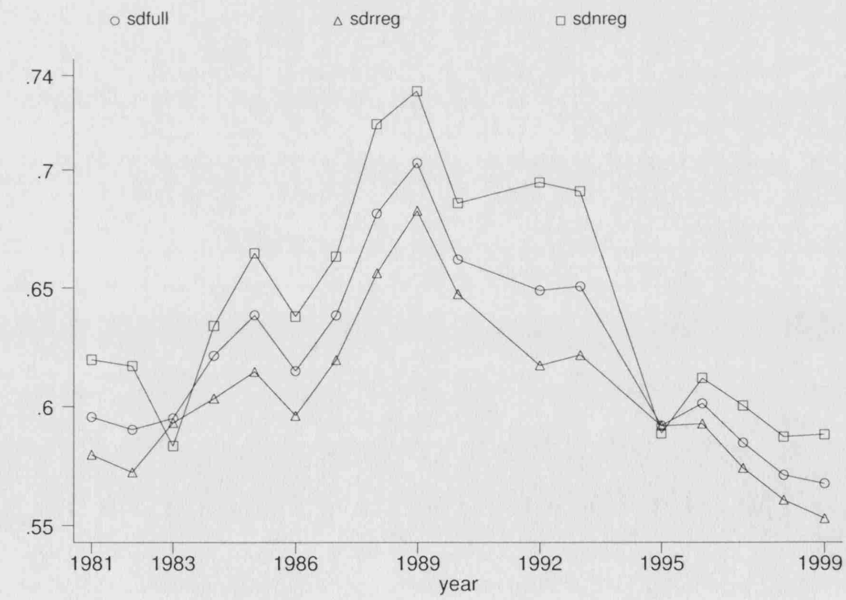

Figure 2.14: Standard Deviation of the Residuals - 1981 - 1999

1981-1999		1990-1999		1992-1999	
V(w) 1981	0.4070	V(w) 1990	0.4647	V(w) 1992	0.4833
V(w) 1999	0.3460	V(w) 1999	0.3460	V(w) 1999	0.3460
Counterfactual	0.3838		0.3518		0.3854
Contribution	62%		5%		29%

Table 2.6: The Effect of the Decline in Wage Differential on the Variance of Log Hourly Wage

following variance decomposition:

$$Var(w) = R * Var(w^R) + (1 - R) * Var(w^{NR}) + R * (1 - R) * (w^R - w^{NR})^2 \quad (2.5)$$

where $Var(w)$ is the variance of the conditional log hourly wages, i.e,. the variance of the residual of the standard wage equation without controlling for the registered status (pooled sample); R is the proportion of registered workers, w^R and w^{NR} are the log hourly wage of registered workers and non-registered workers. It is worth noting that in the decomposition the difference $(w^R - w^{NR})$ was estimated as the value of the dummy coefficient for registered workers in a joint wage equation.

The first two terms in equation (2.5) may be thought of as measuring within-sector changes in the structure of wages, and the third as measuring between-sector changes due to the possession of a work-card.

In order to measure the contribution of the falling wage differential between the two groups, we calculate some counterfactuals assuming that everything else has changed between the chosen baseline years and 1999, but the wage gap has remained constant. The difference between the actual value of the variance and the value given by the counterfactual is a measure of the importance of the falling wage gap for the fall in inequality as whole. Table 2.6 presents these calculations between 1981, 1990, 1992 (baseline years) and the final of the sample period 1999. The counterfactual shows what would have been the variance of the wages if the wage gap were the same as in the baseline year. The difference between the counterfactual and the actual measure, give us an estimate of the contribution of the falling wage gap to the decrease in the variance of log-hourly earnings. The estimates vary widely according to baseline year. But even for a year with a not very high wage gap such as 1990 (see Figure 2.4), the reduction in wage differential was able to explain 5% of the decrease in the variance of earnings. As for the whole period, i.e., from 1981 to 1999, the fall in the wage gap explains 62% of the decrease in the variance of wages, which is quite a high contribution.

Fields (2002) puts forward a methodology designed to account for inequality and to decompose it into the contribution of the explanatory factors[45] of a standard semi-log regression. The decomposition of the log-variance of wage can be written as:

$$s_j(lnw) = \frac{cov[a_j, Z_j]}{\sigma^2(lnw)} = \frac{a_j\sigma(Z_j)cor[Z_j, lnw]}{\sigma(lnw)} \tag{2.6}$$

where s_j is the 'relative factor inequality weight' of the explanatory factor j, a_j is the coefficient of the explanatory factor j in the wage equation[46], Z_j is the explanatory factor j, and $\sigma^2(lnw)$ is the variance of the log wage.

This decomposition allows one to account for the level of wage inequality in a particular country at a particular time, and for a specific group of workers. In order to account for "differences" in inequality over time or between groups, Fields (2002) proposes the following decomposition:

$$\Pi_j(I(.)) = \frac{s_{j,2}I(.)_2 - s_{j,1}I(.)_1}{I(.)_2 - I(.)_1} \tag{2.7}$$

where Π_j is the contribution of the explanatory factor j to the change in inequality as measured by the inequality index I(.) between period 1 and 2[47].

Table 2.7 shows both Fields' decomposition in levels for 1999 and 1981, and the decomposition for changes between these two years. For the full sample the two most important factors explaining wage inequality in both years are the residual and education. The possession of the work-card (registered) is the third more important factor, but its contribution is rather modest when compared to the two other factors. The registered and non-registered samples display a similar pattern, but the residual seems to be more important for the non-registered than for the registered in order to explain the level of inequality.

As for changes in inequality, the first row in Table 2.7 shows that the inequality - as measured by the log variance of hourly real wages - fell for the full sample, and for both registered and non-registered workers. For the full sample the most important factor in reducing inequality was education (0.61), the second most important was the possession of a work-card (0.38), and finally gender (0.28). The other factors played only a minor role in changes in inequality. Regional and the residual changes acted in the opposite direction and would have triggered more inequality. For the sample of registered workers, education is by far the most important factor, whereas for the sample of non-registered workers, the most important factor was gender.

[45]The explanatory factors include all the regressors and the residual of the wage equation.
[46]a_j is equal to 1 when the explanatory factor is the residual of the wage equation.
[47]The index 1 and 2 also can indicate different groups of workers.

	Full Sample			Registered			Non-registered		
	Level		Changes	Level		Changes	Level		Changes
	1999	1981	1999-1981	1999	1981	1999-1981	1999	1981	1999-1981
log variance	0.84	0.97	-0.13	0.78	0.84	-0.06	0.77	0.87	-0.09
Education	0.32	0.36	0.61	0.38	0.44	1.13	0.20	0.19	0.12
Experience	0.04	0.04	0.03	0.03	0.02	-0.17	0.06	0.08	0.27
Region	0.06	0.02	-0.20	0.04	0.01	-0.27	0.08	0.04	-0.30
Metropolitan	0.02	0.02	0.02	0.01	0.01	0.03	0.03	0.02	-0.02
Gender	0.05	0.08	0.28	0.04	0.05	0.19	0.06	0.16	0.99
Registered	**0.07**	**0.11**	**0.38**						
Residual	0.45	0.38	-0.13	0.51	0.48	0.09	0.58	0.51	-0.05
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2.7: Factor Contribution to Wage Inequality and to Change in Inequality: 1981 - 1999

Thus, we can conclude that the possession of work-card was one of the main factors behind the fall in inequality between 1981 and 1999 after education[48].

However, the estimates above offer only a partial view of what happened with the inequality during the sample period. In order to have a better understanding of what happened to inequality at different parts of the wage distribution and to assess the role of observable and unobservable components in shaping the evolution of wage inequality between 1981 and 1999, we will apply Juhn, Murphy and Pierce's (1993) decomposition (full-sample distribution accounting scheme). This decomposition allows one to distinguish which changes in inequality were due to changes in observed quantities (of skills), observed (skill) returns and changes in unobserved returns and quantities (of unobserved skills). The starting point is the estimation of standard earning equations:

$$w_{it} = X_{it}\beta_t + u_{it} \tag{2.8}$$

where w_{it} is the log hourly-wage of individual i in year t, X_{it} is a vector of observed individual characteristics in t, and u_{it} is the log wage residual, which is assumed to be an unknown function of prices and quantities of unobserved skills, measurement error and estimation error. Juhn et al. (1993) assume that the wage equation residual has two components: an individual's percentile in the wage distribution θ_{it} and the distribution function of the residuals $F_t()$, which implies, by the definition of the cumulative distribution function, that one can write the residual as:

$$u_{it} = F_t^{-1}(\theta_{it}|X_{it}) \tag{2.9}$$

where $F_t^{-1}(\theta_{it}|X_{it})$ is the inverse cumulative residual distribution for workers with characteristics X_{it} in year t.

[48]The aggregate result for education is completely due to changes for education level below the complete college group. Changes in the relative factor of the latter were in direction of more inequality.

The decomposition is illustrated by the formula:

$$w_{it} = X_{it}\beta + X_{it}(\beta_t - \beta) + G^{-1}(\theta_{it}|X_{it}) + [F_t^{-1} - G^{-1}(\theta_{it}|X_{it})] \tag{2.10}$$

where β and G^{-1} are the returns to observable skills and the cumulative residual distribution for the base period, respectively. This formula allows one to recover the counterfactual wage distribution implied by holding fixed any subset of the components described above.

In practice, the procedure consists of running wage equations separately for two periods, one of which is the base period, and comparing what would the wage inequality have been if 1) the distribution of individual characteristics ($X's$) in period 2 had remained the same as the distribution in period 1 (base period), holding returns and residuals as in period 2, counterfactual w_{it}^1; 2) both the distribution of individual characteristics and the returns have remained the same as in period 1 and the residuals as in period 2, counterfactual w_{it}^2. Thus, differences in wage inequality between period 1 and 2 can be decomposed in differences due to changes in the observables (changes in w_{it}^1), due to changes in returns (additional changes caused by w_{it}^2), and due to changes in unobservables (changes in w_{it} beyond those found w_{it}^2).

Taking 1981 as the base period we applied the Juhn et al. decomposition to the full sample of employees and separately for registered and non-registered worker samples. As shown in Table 2.8 (Panel A) , the $90-10$ log hourly-wage differential fell for the full sample and both registered and non-registered workers. For the full sample and for registered workers the major contribution for this decrease in inequality came from changes in the returns, whereas for the non-registered workers it came from changes in observables. The residuals also had a positive, but less important contribution. For the full sample and for registered workers, changes in the observable characteristics, unlike the other components, would have contributed to an increase in inequality.

The inequality in the upper part of the wage distribution had different patterns for the different samples. The $90-50$ log wage differential has decreased for the full sample and for registered workers, but has slightly increased for non-registered workers. Again, changes in the observables would have led to a higher wage inequality for the full sample and for registered workers. In the non-registered case, they would have contributed to a fall in inequality, but changes in returns and changes in unobservables were strong enough to compensate its effect and then worsen the wage inequality.

In the lower part of the wage distribution the changes in inequality were different from the ones observed in the upper part. The $50 - 10$ log wage differential has decreased for the full sample and for non-registered workers and increased for registered workers. For the full sample, all three factors acted to deliver a reduction in inequality, but the main effect came from changes in returns. For the non-registered sample change in returns would have contributed to worsen wage inequality, whereas changes in observables and unobservables contributed to attenuate it. As for the registered sample, the small increase in wage inequality was entirely due to changes in observables.

In order to have a better view of what happened during the period we calculated the same decomposition for years with a low wage differential between registered and non-registered workers. As shown above, 1986 (Panel B), 1990 (Panel C) and 1995 (Panel D) were years with a particularly low wage differential. The decompositions for these years – keeping 1981 as the base – reveal some similarities, but also some interesting differences. Among the similarities it is worth noting that in all periods, the $90 - 10$ log wage differential has diminished for the full sample and the main responsible for that was the returns to observable skills. Changes in the observables and in the residual distribution have contributed in the opposite direction, causing more inequality[49]. Another similarity is the fact that inequality has decreased for the upper part of the wage distribution as measured by the $90 - 50$ log wage differential among the registered workers[50] and increased among the non-registered workers. Differently, the inequality in the lower part of the wage distribution, the 50-10 log wage differential has increased among the registered and decreased among the non-registered in all three periods, when compared to the wage structure of 1981.

The dissimilarities between the decomposition for 1981/1999 and the other three periods refer mainly to the fact that the inequality measured by the $90 - 10$ log wage differential has increased for both registered and non-registered workers in 1986 and 1990. For non-registered workers it has decreased in 1995, but for the registered it was still in a higher level than in 1981. Thus, the reduction in inequality observed for the $90 - 10$ log wage differential for the full sample was due to the reduction in the wage differential between the two groups, i.e., to the fact that the non-registered workers have been improving their position within the overall wage distribution. It is worth noticing that the non-registered workers are concentrated in the lower tail of the wage distribution and the inequality as measured by the 50-10 log differential

[49]The exception being 1995, when even the unobservables contributed to the reduction in inequality.

[50]There is an exception for the period 1981-1990. The inequality increased for all measures $90 - 10$, 90-50 and 50-10 log wage differentials for the sample of registered workers.

PANEL A	w_{it}=1981	w_{it}=1999	w_{it}^1	w_{it}^2	Δobservables	Δreturns	Δunobservables
90-10							
overall	2.435	2.058	2.483	2.081	-12.7%	106.6%	6.1%
registered	2.071	1.990	2.205	2.032	-165.6%	213.8%	51.8%
non-registered	2.186	1.903	1.816	2.109	130.9%	-103.6%	72.8%
90-50							
overall	1.360	1.195	1.411	1.203	-31.1%	126.1%	4.9%
registered	1.318	1.206	1.342	1.222	-20.6%	106.4%	14.3%
non-registered	1.071	1.088	0.918	1.055	-923.7%	824.5%	199.2%
50-10							
overall	1.075	0.863	1.072	0.877	1.5%	91.5%	7.0%
registered	0.753	0.784	0.864	0.810	350.7%	-168.7%	-81.9%
non-registered	1.115	0.815	0.898	1.054	72.3%	-52.1%	79.8%
PANEL B	w_{it}=1986	w_{it}=1999	w_{it}^1	w_{it}^2	Δobservables	Δreturns	Δunobservables
90-10							
overall	2.435	2.280	2.493	2.251	-37.8%	156.7%	-18.9%
registered	2.071	2.159	2.156	2.119	96.8%	-42.2%	45.4%
non-registered	2.186	2.188	2.167	2.200	-850.0%	1466.6%	-516.6%
90-50							
overall	1.360	1.355	1.406	1.328	-995.2%	1667.9%	-572.8%
registered	1.318	1.303	1.339	1.287	-133.6%	337.7%	-104.1%
non-registered	1.071	1.183	1.068	1.088	-3.2%	18.2%	85.0%
50-10							
overall	1.075	0.925	1.088	0.923	-8.2%	110.0%	-1.8%
registered	0.753	0.856	0.818	0.832	63.2%	13.2%	23.6%
non-registered	1.115	1.006	1.099	1.112	14.4%	-11.8%	97.4%
PANEL C	w_{it}=1990	w_{it}=1999	w_{it}^1	w_{it}^2	Δobservables	Δreturns	Δunobservables
90-10							
overall	2.435	2.430	2.528	2.321	-1909.0%	4234.7%	-2225.7%
registered	2.071	2.363	2.244	2.210	59.2%	-11.4%	52.2%
non-registered	2.186	2.252	2.256	2.132	106.4%	-188.7%	182.3%
90-50							
overall	1.360	1.417	1.442	1.361	144.1%	-141.8%	97.8%
registered	1.318	1.376	1.351	1.310	57.1%	-72.1%	115.0%
non-registered	1.071	1.279	1.246	1.185	83.9%	-29.1%	45.2%
50-10							
overall	1.075	1.013	1.086	0.960	-17.6%	202.8%	-85.2%
registered	0.753	0.986	0.892	0.901	59.8%	3.7%	36.5%
non-registered	1.115	0.973	1.010	0.947	73.5%	44.9%	-18.4%
PANEL D	w_{it}=1995	w_{it}=1999	w_{it}^1	w_{it}^2	Δobservables	Δreturns	Δunobservables
90-10							
overall	2.435	2.195	2.529	2.198	-39.1%	137.6%	1.4%
registered	2.071	2.163	2.223	2.129	165.7%	-102.8%	37.1%
non-registered	2.186	1.911	2.184	1.950	0.7%	85.0%	14.3%
90-50							
overall	1.360	1.320	1.447	1.302	-222.9%	368.5%	-45.7%
registered	1.318	1.288	1.360	1.261	-134.8%	322.6%	-87.8%
non-registered	1.071	1.128	1.245	1.132	308.4%	-200.6%	-7.8%
50-10							
overall	1.075	0.874	1.081	0.896	-3.1%	92.4%	10.6%
registered	0.753	0.875	0.863	0.868	90.6%	3.6%	5.9%
non-registered	1.115	0.783	0.939	0.818	53.0%	36.5%	10.5%

Table 2.8: Juhn-Murphy-Pierce Changes in Inequality Decomposition

decreased in 1986, 1990, 1995 and 1999. Besides, it was the reduction in lower tail of the wage distribution for the full sample that led to the reduction in the $90-10$ log differential between the four years analysed and 1981, as noticed above.

One possibility that could explain these results, mainly from 1995 onwards is the substantial reduction of non-registered workers earning less than the minimum wage. Besides, the surprising increase in the indexation of non-registered earnings to the minimum wage may have contributed to a decrease in the wage inequality among low wage non-registered workers[51].

2.3 Concluding Remarks

Most of the empirical literature on segmentation between formal and informal sectors in developing countries focuses on the wage differential between these two groups. As shown above, a major problem in comparing these studies is the lack of homogeneity in the way the informal sector is defined. Another characteristic of the empirical literature is the great emphasis on the need to correct for selectivity bias when discussing the hypothesis of segmentation in the lines of formal/informal sector[52]. In the next chapter we will discuss selectivity issues in the context of a job queue for formal jobs. We will assess among other things, the role of the wage differential in determining sector allocation and evaluate how the results change once we take into account the individual "willingness" to switch from a informal to a formal job[53].

As documented above there were great changes in the proportion of registered/non-registered workers in the economy, in its wage differential and in the inequality within and between the two groups over the 1980's and the 1990's. According to Fields' (2002) decomposition scheme changes in the variable related to the possession of a work-card was the most important force, after education, driving down the variance of the log hourly real wage. The most striking and lasting changes in the wage gap

[51]In Chapter 5 we will discuss in depth the indexation of the wage of informal and formal sector to the minimum wage.

[52]Due to the difficulties in finding a common set of exclusion restrictions that could be used throughout the period analysed we did not control for selectivity in this chapter. However, ou results in chapter 3 as well as the literature results in Tanuri-Pianto and Pianto (2002) and in Carneiro and Henley (2001) show that the estimated coefficients of the selectivity terms do not change much as well as the $R2$ of the wage equation estimations - notice that the estimations of these studies are far apart in time: from 1990 until 1999. In this sense it would seem quite unlikely to me that the addition of the IMR coefficients would change the results of the decompositions analyses carried out in this chapter.

[53]Unfortunately we can only test the job queue hypothesis with the 1990 data. It would have been interesting to analyse whether or not the queue has diminished after the trade liberalisation reform, specially, with the fall in the wage gap between formal and informal workers.

occurred after 1990. Episodical changes such as the increase in the proportion of non-registered workers in 1983 or the sharp reduction in the wage gap in 1986 did not last. In contrast, the increase in the proportion of non-registered workers observed since 1990 and the diminishing wage gap after 1992 seem to be a more stable process triggered by recent moves in the Brazilian economy. In Chapter 4, we will investigate whether the fall in the wage gap between registered and non-registered workers and the increase in the size of the informal sector are correlated with the process of trade liberalisation of the early 1990's. Similarly, in Chapter 5 we will investigate whether minimum wage hikes had any effect on the employment mobility for workers in the formal and in the informal sectors. Particular attention will be given to the growing indexation of the informal sector wage to the minimum wage after 1990, as documented in this chapter.

Chapter 3

Do Informal Workers Queue for Formal Jobs in Brazil?

3.1 Introduction

As seen in Chapter 2 there is a large wage differential between formal and informal sector workers in Brazil. This differential may be even larger if one takes into account the total compensation package of a formal job contract. This package includes mandatory annual bonus (usually one extra monthly wage per year), paid holidays and access to public funds such as unemployment benefit and severance payment. The form of financing these extra-benefits may put some extra burden on employers. Social security contribution in one of the main non-wage costs that employers have to pay, but besides that there is also the severance payment that includes a contribution to a compulsory fund and a fine equal to 40% of the amount of money in the compulsory fund. Moreover, there are same payroll tax that finance other social funds. It is easy to see how employers may prefer to avoid registration, but some commentators also argue that employees may find it profitable to avoid registration. This would be the case for workers for whom the social security pension would be very close to the minimum wage - since they could get that pension out of the social assistance system anyway - and/or for those who severance payment would never grow enough to be a real compensating package.

However, as discussed in the review of the literature on segmentation in the lines of the formal/informal sector, the wage gap between (observably) similar workers allocated into these two sectors is not enough to prove the existence of segmentation in the labour market[1].

[1] One can argue that his/her definition of segmentation consists only of the existence of wage differential between otherwise similar workers employed in different sectors, but the existence of

Maloney (1997) argues that the wage differential is not a good guide to assess the existence of segmentation and claims that tests on the existence of a job queue is a much clearer indicator of segmentation. If some workers would prefer to get a formal job, but do not get it and at the same time similar workers do get it, then one could claim the existence of some sort of segmentation. This evidence could be an indication of rigidity in the labour market that can be caused by the excess of regulation or simply by efficiency wage practices or a combination of both of them[2]. The existence of a job queue for formal jobs, therefore, has implications for both economic efficiency and long-run income inequality. As for the first aspect, law reform making the labour market more flexible has been listed as one of the important reforms that must be implemented by developing and developed countries that face slow growth and increasing unemployment or under-employment[3]. The existence of job queue can be considered an indicator of this lack of flexibility. As for income inequality, the major problem is that formal jobs come together with several benefits, including pensions, that are not readily available for informal sector workers. The longer the individual spends in the queue for formal jobs, the more likely is that he/she will lose these benefits or get a smaller proportion of them; perpetuating and increasing inequality. Differently, it has been argued that the possibility of tax evasion and more flexible hours were factors that would compensate for the lack of fringe benefits and access to social funds associated with the formal sector jobs (Maloney, 1997). However, if this were the case, then one should observe a high proportion of informal sector workers satisfied with their current occupational status. According to the 1990 PNAD, this can be considered the case for self-employed workers, but hardly can be considered consistent with the view that non-registered wage workers have of their own current status since 70% of them would like to switch to a formal job.

Most of the studies on labour market segmentation in developing countries have focused on the need to correct the wage equations for both formal and informal sectors in order to get unbiased estimates of the wage differential between the two sectors[4]. This procedure, however, does not limit itself to finding the unbiased wage differential between the formal and informal sectors, i.e, the differences in the wages that any individual drawn from the population would get in either sector.

wage differential between similar workers is also compatible with other competitive explanations such as compensation differential, comparative advantages or returns to entrepreneurial ability (in the case of self-employment).

[2]See review in Section 2.1 in Chapter 2

[3]See Heckman and Pages(2004) for a defence of labour market reform as a way to improve job creation in Latin American and Caribbean countries.

[4]See Table 2.1 in Chapter 2 for an overview of the these studies.

It also allows one to test whether or not individuals select themselves into these two sectors according to their comparative advantages. The idea that individuals self-select into the occupation/sector where they are more productive dates back to Roy's (1951) model[5]. Roy shows how individuals choose to work in the occupation in which they are more productive and how the flow and the stock of workers in each occupation varies with demand shocks in the product of their industry. One of the key assumptions of the Roy model is that there are no barriers to entry in any of the sectors. Thus, there is no job queue in such framework.

It is interesting to note that despite all the controversy around the existence of rationing in the formal sector and the discussions on how to correct for selectivity bias, the literature on job queuing has been scarcely applied to tackle this issue[6]. This literature, however, has been widely applied in the study of union/non-union wage differential, private and public wage differential, and queue for minimum wage jobs[7].

According to the job queue approach the process of sector allocation cannot be correctly modelled by a univariate process that determines whether the individual prefer to work in one sector or other based on his/her comparative advantage and/or preferences. If jobs are rationed, one must take into account the employers' criteria for choosing workers for rationed jobs. Individual characteristics and past employment history, for instance, may have different effects on the probability of desiring a formal job (in the queue status) and on the probability of being chosen (from the queue) by the employer for that job. In addition, the existence of a job queue for formal jobs would lead to a biased estimation of the wage equation, since the univariate specification of the selection equation would misrepresent the process of sector allocation (Farber, 1983). Abowd and Farber (1982) and Farber (1983) were the first to apply this idea to the allocation of workers into union/non-union sectors, but they did not go further to correct the wage equation using the double selectivity criteria. Mengistae (1998), following Maddala's (1983) suggestion, puts forward a methodology that not only incorporates the double selectivity criteria in the wage equation, but also enables the estimation of the proportion of workers

[5]The major difference between Roy's model and other models of comparative advantage in the choice of sector/occupation is that in Roy's model only the incomes are compared in the decision rule, whereas in more general models of utility maximization, non-wage dimensions also play a role in the sector choice, see Heckman and Honore (1990).

[6]We are only aware of Maloney (1997) attempt to test the existence of job queue for formal jobs in Mexico in a time series context.

[7]See Abowd and Farber (1982) and Farber (1983) for union/non-union, Venti(1987), Heywood and Mohanty(1995), Mohanty(1998) and Mengistae(1998) for private/public sector and Holzer, Katz and Krueger (1991) for minimum wage applications.

queuing directly from the corrected wage equation for workers who were not chosen from the queue. The main weakness of this approach consists in the fact that, in general, one has to rely on the estimation of bivariate probit with severe partial observability, since the "in the queue" status is hardly observed. In such a situation the identification of the two selection equations relies heavily on nonlinearities in the functional form of the probability distribution (Farber, 1983). More worryingly, procedures based on Poirier's (1980) bivariate probit with partial observability are known to have a very bad performance in terms of convergence[8]. Thus, attempts to add more information on the severe partial observability of the "in the queue" status may help to identify the two equations[9].

We will use in this chapter a special supplement of 1990 Brazilian Annual Household survey (PNAD) to identify workers in the informal sector who would be queuing for formal jobs in order to lessen the severe partial observability of Abowd and Faber (1982) and Poirier (1980) bivariate probit models. Because of this we will be able to identify the two equations separately without relying heavily on their functional form and on the exclusions restrictions we have to impose to identify separately the two equations. In sum, we are getting from these survey questions some information that the model would help us to guess through very strong assumptions.

The chapter is structured as follows. First, we test the existence of a job queue for formal jobs and how its "length" varies for different groups of workers. This is important for public policy to target more vulnerable workers if one assumes that an informal sector job is a "second best" option. Second, we estimate selectivity-corrected wage equations for formal and informal workers in order to assess the role of the wage differential in determining whether or not a worker join the queue for a formal job (endogenous switching regression). Finally, we investigate the sensitivity of job queue estimates to different assumptions regarding the sector allocation process. Using subjective survey questions on whether or not an informal worker would accept a formal job we are able to estimate a bivariate probit with sample selection[10] and partially overcome the difficulties imposed by the severe partial observability on both Abowd and Faber (1982) and Poirier (1980) models.

[8]It is common to find papers that exclude some variables of the estimation process or redefine them only to guarantee convergence. See, for instance, Mohanty (1998).

[9]Farber (1983) uses the intention to vote in favour of unionisation in order to help the identification of the workers who would be in the queue.

[10]This sort of bivariate probit is also known as Meng and Schmidt bivariate probit or censored bivariate probit.

3.2 Endogenous Switching Regression in the Presence of Job Queue

Most endogenous switching regression models are based on a univariate probit (or logit) process that models the worker's preference for a specific type of job. The assumption behind this approach is that once a worker decides to get a job in a specific sector there is nothing preventing him/her to get hired. Therefore the probability of desiring a formal job is identical to the probability of having a job in that sector. Besides it is assumed that workers base their decision taking into account comparative advantages they would have in the chosen sector. Thus, the wage differential between what a worker would get in a sector *vis-à-vis* what he/she would get in the other sector should play a substantial role in determining the actual sector allocation. The common procedure in this case is to estimate in a first step a reduced form probit (or logit), then correct the wage equations for selectivity bias (due to non-random selection into the sector), get the wage differential based on the corrected wage equations and, finally, estimate a structural probit that incorporates the wage differential as an additional regressor.

To assess the existence of a job queue in an endogenous switching regression model we will relax the assumption that the probability of getting a formal job is equal to the probability of willing to be in the formal sector. The probability of being in the formal sector is the result of two independent decision processes: the decision to join the queue for formal jobs by the worker and the decision to hire a worker who is in the queue by the formal employers. A worker is only observed in the formal sector if he/she both had joined the queue and was chosen from the queue.

A worker decides to join the queue if the utility that he/she derives from this choice is higher than a specific threshold. In the case of the formal sector, he will choose this sector if the "advantages" related to a formal job (e.g. pensions, paid vacations) more than compensate possible shortcomings related to it (e.g. higher difficulty in evading income taxes, longer working hours). This decision can be modelled by the latent variable I_{1i}^* that summarizes the willingness to get a formal job:

$$d_{1i} = \begin{cases} 1 & \text{if } I_{1i}^* = X_{1i}'\beta_1 + u_{1i} > 0 \\ 0 & \text{if } I_{1i}^* = X_{1i}'\beta_1 + u_{1i} \leq 0. \end{cases} \tag{3.1}$$

where d_{1i} is an indicator equal to one if the worker is in the queue and equal to zero if the worker is not in the queue, X'_{1i} is a vector of individual characteristics assumed to determine the individual decision to join the queue, β_1 is a vector of parameters and u_{1i} is an idiosyncratic component that captures unobserved heterogeneity in the preference for a formal job and omitted variables.

Similarly, the employer's decision can be modelled by a latent variable, I^*_{2i}, that captures his/her perception about the worker's productivity. Formal jobs are associated to higher firing and hiring costs and also to some mandatory fringe benefits. When deciding to hire a worker, a formal employer must evaluate whether or not the worker's productivity compensates the overall cost. The employer's decision, then, is based on the difference between the worker's productivity and his/her associated costs.

$$d_{2i} = \begin{cases} 1 & \text{if } I^*_{2i} = X'_{2i}\beta_2 + u_{2i} > 0 \\ 0 & \text{if } I^*_{2i} = X'_2\beta_2 + u_{2i} \leq 0. \end{cases} \tag{3.2}$$

where d_{2i} is an indicator equal to one if the worker is chosen from the queue and equal to zero if the worker is not chosen from the queue, X'_{2i} is a vector of observable individual characteristics assumed to determine the employer's decision to select a worker from the queue, β_2 is a vector of parameters and u_{2i} is an idiosyncratic component that captures unobserved heterogeneity in the employer's perception and/or omitted variables.

The endogeneity of the switching model resides in the fact that the difference between the worker's predicted wage in the formal sector and the worker's predicted wage in the informal sector enters the "in the queue" equation as an additional explanatory variable. Hence, sector allocation determines wage and wage differential determines sector allocation. The first problem to overcome in order to estimate a structural probit (or bivariate probit) like that is to get the correct equation to predict the wage differential. The wage equations for the formal and informal sectors can be estimated as:

$$W_{fi} = Z'_{fi}\gamma_f + \epsilon_{fi} \tag{3.3}$$

$$W_{ii} = Z'_{ii}\gamma_i + \epsilon_{ii} \tag{3.4}$$

where W_{fi} and W_{ii} are the log hourly wage rate for formal sector and informal workers respectively, Z_{fi} and Z_{ii} are the variables that determine wages including both individual and industry characteristics and γ_f and γ_i are vectors of parameters. However, it is well known that OLS estimates can be biased for not taking into account the sector allocation decision made by workers, and in a job queue approach also by the employers. OLS estimates assume that $E(\epsilon_{fi}|X_{1i}) = 0$, but W_{fi} is only observed if the worker queued for a formal job and was chosen from the queue. This condition, in general, may lead to the violation of the assumption that the conditional expectation of the residuals is equal to zero, since the sample of formal workers is not randomly drawn from the population. The violation is due to the fact that there would be correlation between Z_{fi} and ϵ_{fi} operating through the relationship between ϵ_{fi} and the pair u_{1i} and u_{2i} in equations (3.1) and (3.2), respectively. More specifically estimates of equation (3.3) will be biased if $E(\epsilon_{fi}|X_{1i}, I_{1i}^* > 0 \quad \text{and} \quad I_{2i}^* > 0) \neq 0$.

The wage equation for informal workers is also censored, but the rule is a bit more complicated than the one for the formal sector. This is due to the fact that there are, at least, two different types of informal workers. The ones who are in the queue and were not chosen from the queue – $E(\epsilon_{ii}|X_{1i}, I_{1i}^* > 0 \quad \text{and} \quad I_{2i}^* < 0)$ - and the ones who did not join the queue – $E(\epsilon_{ii}|X_{1i}, I_{1i}^* < 0)$. In the next section we will discuss this point more thoroughly since its estimation is dependent on the assumptions we make regarding the degree of observability of the "in the queue" status[11].

Similarly to the "in the queue" equation, the wage that a worker would command should he/she be hired by a formal employer may also enter as an extra explanatory variable in the "chosen from the queue" equation. As mentioned earlier, formal employers try to minimize labour cost conditional on worker's productivity. Hence, given the productive characteristics of a worker, the wage he/she would get in the formal sector should enter the "chosen from the queue" with a negative sign. The more expensive the worker given his/her productivity potential, the lower the probability that he/she would be selected from the queue. Again, we face the task of estimating the wage that any worker in the population would get once in the formal sector.

In this context a worker would be found in the formal sector if:

$$I_{1i}^* = \alpha_1(W_{fi} - W_{ii}) + X_{1i}'\beta_1 + u_{1i} > 0$$

and

$$I_{2i}^* = \alpha_2(E(W_{fi}|I_{1i}^* > 0)) + X_{2i}'\beta_2 + u_{2i} > 0 \tag{3.5}$$

[11]If we had full observability we should also observe the ones who were not in the queue and would be chosen from it.

where $E(W_{fi}|I_{1i}^* > 0)$ is the expected wage that a worker would get in the formal sector should he/she be in the queue; α_1 is the parameter of the wage differential and α_2 is the parameter of the predicted formal wage should the worker be in the queue. X'_{1i}, X'_{2i}, u_{1i}, u_{2i}, β_1, and β_2 are defined as before[12]. In the next section we discuss how to estimate this structural bivariate probit according to different assumptions about the degree of observability of the "in the queue" status.

3.3 Estimation Procedure

Endogenous switching regression models such as the one described in the previous section can be estimated either in two or three steps. The two-step procedure would imply the joint estimation of the "in the queue" and "chosen from the queue" reduced form equations and of the wage equations for the two sectors - formal and informal - through the maximization of a likelihood function and then the estimation of a structural bivariate probit with the predicted wage differential as an additional regressor.

The three-step procedure would imply the estimation of a reduced form bivariate probit whose residuals would be used to approximate the non-zero expectation of ϵ_{fi} and ϵ_{ii} conditional on u_{1i} and u_{2i}. These approximations, the so-called Inverse Mills ratio, would enter the wage regression for each sector in a second step in order to restore the assumption of zero expectation of the wage regression residuals. Finally, the structural bivariate probit would be estimated using as additional regressors both the (offered) wage differences estimated from the wage regression of the second step in the "in the queue equation" and the wage that a worker in the formal sector would get should he/she be in the queue in the "chosen from the queue" equation.

A major issue that affects both the maximum likelihood (two-step procedure) and the three-step procedure refers to their reliance on the trivariate normality of the residuals of the sector allocation equation and of the wage equations. The assumption of normality makes the correct specification of the model more important than in the usual linear regression framework[13]. Recent research has relaxed this assumption by estimating the sector allocation equation without assuming any particular distribution, but the conclusions of these semiparametric attempts to correct for selectivity in order to overcome the bias triggered by erroneously assuming a

[12]In the next section we will address the issue of identification, when it will become clear that X'_{1i}, X'_{2i} in the structural bivariate probit are in fact a subset of X'_{1i}, X'_{2i} used in a first step in order to estimate the reduced form of the bivariate probit.

[13]This assumption is particular strong when we know that there is some mass concentration on the minimum wage and its multiples.

trivariate (or bivariate) normal distribution for the residuals are mixed at best[14].

Another problem related to the strategies to correct for selectivity bias is the issue of identification. The need for an exclusion restriction in order to separately identify the selection and wage equations is not relaxed by semiparametric techniques; on the contrary, it makes it even more restrictive, since it requires that at least one of the excluded variables to be continuous (Heckman, 1990). Therefore, one has to find a variable that determines the "in the queue" status, but does not determine the "chosen from the queue" status (or vice-versa). Moreover an additional exclusion restriction is necessary in order to identify the parameters of the wage equations. So, we have to find, at least, one variable that determines the wage, but does not affect either the "in the queue" or "the chosen from the queue" decisions[15].

So far we have skipped the discussion of the implication of partial observability for the estimation of both reduced and structural forms bivariate probit of the job queue model. In general, survey data only bring information on the actual status of workers: formal or informal. We do not know whether informal workers wish to get a formal job or if they prefer to stay as informal workers for whatever reason. Thus, we do not observe either d_{1i} or d_{2i}, but only the product $d = d_{1i} * d_{2i}$. If this product is equal to one, the worker is in the formal sector, if it is equal to zero he/she is in the informal sector. Note that besides the lack of information on whether informal workers are in the queue for formal jobs, we also have no information on whether or not formal employers would like to hire workers who are not in the queue. This latter type of partial observability is modelled by Poirier (1980) and assumes that both decisions - join the queue and choose from the queue - are taken simultaneously. Abowd and Farber (1982) assume that the sector allocation process can be modelled as a bivariate probit with partial observability and sequential decision as described in the last section. Employers only hire workers from the pool of workers who are in the queue, thus the distribution of u_{2i} is defined only over the subpopulation for which $d_{1i} = 1$. In this case, one can only make conditional inferences, whereas in Poirier's procedure allows both conditional and marginal inferences (Maddala, 1983).

A shortcoming of the bivariate probit with partial observability discussed above is the severe degree of partial observability. The dependent variable in both the

[14]See for instance Newey, Powell and Walker (1990) for an application of semiparametric methods to correct for selectivity that do not imply major differences in relation to the standard "normal-based" procedure. However, Lanot and Walker (1998), Schafgans (1998) and Martins (2001) find substantial differences in the results yielded by those two procedures.

[15]Note that the reduced form bivariate probit includes the variables used to identify the wage equation, nevertheless, they are excluded from both the "in the queue" equation and the "chosen from the queue equation" in the estimation of the structural bivariate probit.

"in the queue" equation and the "chosen from the queue" equation is the same. This occurs because we do not observe the two types of informal workers: the ones who are not in the queue and the ones who are in the queue and are not chosen from it. The Abowd and Farber approach aims to fulfill this lack of information by matching the characteristics of workers for which $d = 1$ with those of workers for which $d = 0$, after having distinguished the set of characteristics that determine the probability that $I_1^* > 0$ through their effect on the "in the queue" status from the set of characteristics that influence the same probability through their effect on the "chosen from the queue" status. Therefore, the exclusion restrictions play a fundamental role on this identification process.

One way to get more information is to rely on survey questions that try to measure the willingness to get a formal job[16]. If we get an information like that, e.g., if the informal worker would like to switch to a formal job, we could estimate a bivariate probit with sample selection[17]. It is important to note that despite having two different dependent variables (defined over different populations), we still do not have full observability. This is so because we do not observe formal workers who would like to become informal workers. However, if we assume that there are no barriers to entry in the informal sector, we can say that all formal workers desire a formal job. Another feature of this approach is that similar to Abowd and Farber (1982) approach, it assumes a sequential decision process, implying that formal employers only hire workers from the pool in the queue.

Assuming that u_{1i} and u_{2i} have a bivariate normal distribution with means zero and unit variances and zero covariance[18], the likelihood function to be maximized for Abowd-Farber bivariate probit with partial observability and sequential decision process can be written as[19]:

$$L_1 = \prod_{d=1}[\Phi(\tilde{X}_{1i}'\beta_1)\Phi(\tilde{X}_{2i}'\beta_2)]\prod_{d=0}[1 - \Phi(\tilde{X}_{1i}'\beta_1)\Phi(\tilde{X}_{2i}'\beta_2)] \tag{3.6}$$

[16]Farber (1983) adopts a similar strategy and adds the information of whether or not the non-union worker would vote in favor of unionization.

[17]This approach is quite common in attempts to control for both labour market participation and employment decision. As we only observe the employment status of the workers who participate in the labour market, the bivariate probit is censored for the sample that does not participate in the labour market. See Mohanty(2002) and Co et al. (2002) for interesting applications.

[18]Note that the the Abowd-Farber bivariate probit assumes that both decisions are uncorrelated.

[19]Despite being called Abowd and Faber bivariate probit, this approach was not empirically implemented in their 1982 paper. Actually, in that paper they estimate a modified likelihood function that incorporates the concept of "job rights". They combine a simple probit for the ones with job right and a bivariate probit for those without "job rights". Workers with "job rights" are the ones who held a union job in the previous year, were not fired from that job and did no quit that job to take another union job.

where $d = d_{1i} * d_{2i}$, i.e, $d = 1$ is equal to the set of workers observed in the formal sector and Φ is the normal cumulative distribution and $\tilde{X}_{ji}$ are the explanatory variables of the in the queue ($j = 1$) and the chosen from the queue ($j = 2$) equations, including the variables that are assumed to affect only wages, since this first step corresponds to the estimation of a reduced form bivariate probit.

If one assumes that the sector allocation process is based on a joint (simultaneous) decision as put forward by Poirier (1980), the log likelihood function to be maximized can be written as:

$$L_2 = \prod_{d=1} \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho) \prod_{d=0} [1 - \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)] \tag{3.7}$$

where where $d = d_{1i} * d_{2i}$, i.e, $d = 1$ is equal to the set of workers observed in the formal sector and Φ_2 is the bivariate normal cumulative distribution and ρ is the covariance between u_{1i} and u_{2i}[20].

Note that if we assume that there is no correlation between the residual of the two equations, $\rho = 0$ in equation (3.7), the joint decision model becomes identical to the sequential decision model in equation (3.6)[21].

Finally, the likelihood function of the bivariate probit with sample selection for the in the queue and the chosen from the queue equations is:

$$L_3 = \prod_{d_1=1,d_2=1} \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho) \prod_{d_1=1,d_2=0} \Phi_2(\tilde{X}'_{1i}\beta_1, -\tilde{X}'_{2i}\beta_2; -\rho) \prod_{d_1=0} \Phi(-\tilde{X}'_{1i}\beta_1) \tag{3.8}$$

It is worth noting that in this latter case we clearly have two different dependent variables, one for each "selection" equation. This is so because we observe separately $d_{1i} = 1$, workers in the queue: those who are in the informal sector and who would like to switch to the formal plus those in the formal sector, and $d_{2i} = 1$ those who are in the formal sector, and therefore were chosen from the queue.

A first test for the existence of a job queue for formal jobs can be implemented through the imposition of restrictions on the bivariate estimation. Abowd and Farber (1982) suggest that despite the univariate probit not being nested in the bivariate probit, the adequacy of the bivariate specification can be assessed through a Likelihood Ratio test that constrains all the coefficients of the chosen from the queue equation to zero ($\beta_2 = 0$), with exception of the constant term that is arbitrarily

[20]This is so because $Var(u_{1i}) = 1$ and $Var(u_{2i}) = 1$ due to normalization.

[21]Maddala (1983) argues that if the aim of the research were to find out which factors influence the employer's decision in hiring a specific type of worker, a simultaneous framework would be more appropriate.

fixed as a positive number large enough to ensure that all workers in the queue are chosen from the queue, i.e., $P(I_2^* > 0 | I_1^* > 0) = 1$. Mengistae (1998) extends this idea and puts forward the same test for the "in the queue" equation, so that one can test the hypothesis of a universal queue for formal jobs, i.e., there is no worker who would prefer to stay in a informal job. The procedure is symmetrical to the one described above. All coefficients of the "in the queue" equation are constrained to zero, with exception of the constant that is fixed as a positive number large enough to ensure that all workers are in the queue for formal jobs.

Based on the results of the bivariate probit, one can estimate the probability that workers with a specific set of characteristics are a) found in the formal sector; b) in the queue and c) chosen from the queue conditional on being in the queue[22]. In the sequential case with partial observability, these probabilities can be calculated, respectively, as:

a1) $Prob(I_{1i}^* > 0 \text{ and } I_{2i}^* > 0) = \Phi(\tilde{X}_{1i}'\beta_1)\Phi(\tilde{X}_{2i}'\beta_2)$

b1) $Prob(I_{1i}^* > 0) = \Phi(\tilde{X}_{1i}'\beta_1)$

c1) $Prob(I_{2i}^* > 0 | I_{1i}^* > 0) = \Phi(\tilde{X}_{2i}'\beta_2)$

For the simultaneous case with partial observability and for the sample selection case we would have[23]:

a2) $Prob(I_{1i}^* > 0 \text{ and } I_{2i}^* > 0) = \Phi_2(\tilde{X}_{1i}'\beta_1, \tilde{X}_{2i}'\beta_2; \rho)$

b2) $Prob(I_{1i}^* > 0) = \Phi(\tilde{X}_{1i}'\beta_1)$

c2) $Prob(I_{2i}^* > 0 | I_{1i}^* > 0) = \Phi_2(\tilde{X}_{1i}'\beta_1, \tilde{X}_{2i}'\beta_2; \rho)/\Phi(\tilde{X}_{1i}'\beta_1)$

An estimate of the length of the queue, q, can be obtained by the inverse of the average probability of being chosen from the queue given that the worker joined the queue, $Prob(I_{2i}^* > 0 | I_{1i}^* > 0)$ (Farber, 1983 and Venti, 1987) for each one of the bivariate models:

$$q_1 = \frac{N}{\sum_{i=1}^{N} \Phi(\tilde{X}_{1i}'\beta_1)} \tag{3.9}$$

and

$$q_2 = \frac{N}{\sum_{i=1}^{N} \Phi_2(\tilde{X}_{1i}'\beta_1, \tilde{X}_{2i}'\beta_2; \rho)/\Phi(\tilde{X}_{1i}'\beta_1)} \tag{3.10}$$

where N is the number of observations.

As seen in the last section, the OLS estimates of the wage equation for both

[22]One could also estimate the probability that a worker would be chosen from the queue, but this probability is only meaningful in the simultaneous case, since as discussed above, only in this case it is possible to make marginal inferences regarding the chosen from the queue equation.

[23]This feature does not imply that both methods yield the same results since their likelihood functions are different.

formal and informal workers can be biased due to the censoring rules generated by the sector allocation process. Heckman (1979) puts forward a technique to approximate the non-zero expectation of the residuals in the univariate (probit) case. The extension for the bivariate case is quite straightforward. Assuming that the residual has a normal distribution the first moment of a truncated distribution from below is given by (Maddala, 1983):

$$E(\epsilon_{fi}|I_1^* > 0, I_2^* > 0) = \sigma_{1f}\lambda_{1fi} + \sigma_{2f}\lambda_{2fi} \tag{3.11}$$

where

$$\lambda_{1fi} = \frac{\phi(\tilde{X}'_{1i}\beta_1)\Phi(\tilde{X}'_{2i}\beta_2 - \rho\tilde{X}'_{1i}\beta_1)/\sqrt{(1-\rho^2)}}{\Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)} \tag{3.12}$$

and

$$\lambda_{2fi} = \frac{\phi(\tilde{X}'_{2i}\beta_2)\Phi(\tilde{X}'_{1i}\beta_1 - \rho\tilde{X}'_{2i}\beta_2)/\sqrt{(1-\rho^2)}}{\Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)} \tag{3.13}$$

and σ_{1f} and σ_{2f} are, respectively, the correlation between the residuals of the wage equation of the formal sector and the residuals of the "in the queue" equation and of the "chosen from the queue" equation. Then we can approximate the non-zero expectation by the inclusion of estimates of λ_{1fi} and λ_{2fi}, the so-called Inverse Mills ratios, in the OLS regression. This procedure yields consistent estimates of the wage equation parameters and also allows us to test for the presence of selectivity in the formal sector via a simple t-test of significance of the parameters σ_{1f} and σ_{2f}[24].

However, the bivariate probit with partial observability and sequential decision implies that the correlation between the "in the queue" equation and "the chosen from the queue" equation is (assumed) to be zero $\rho = 0$. In this case (3.12) and (3.13) simplify to:

$$\lambda_{1fi} = \frac{\phi(\tilde{X}'_{1i}\beta_1)}{\Phi(\tilde{X}'_{1i}\beta_1)} \tag{3.14}$$

and

$$\lambda_{2fi} = \frac{\phi(\tilde{X}'_{2i}\beta_2)}{\Phi(\tilde{X}'_{2i}\beta_2)} \tag{3.15}$$

Therefore the wage equation for workers in the formal sector can be estimated as:

$$E(W_{fi}|d = 1) = Z'_{fi}\gamma_f + \sigma_{1f}\lambda_{1fi} + \sigma_{2f}\lambda_{2fi} + \eta_{fi} \tag{3.16}$$

where $\eta_{fi} = \epsilon_{fi} - \sigma_{1f}\lambda_{1fi} - \sigma_{2f}\lambda_{2fi}$ so that $E(\eta_{fi}|d = 1) = 0$.

The unknown Inverse Mills ratio λ_{1fi} and λ_{2fi} can be approximated by substi-

[24] Ahn(1992) shows that the F-statistic of a model with double selection criteria is asymptotically equivalent to the LM test statistic.

tuting the bivariate probit estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ in (3.12) and (3.13) or (3.14) and (3.15) according to the model in use, simultaneous or sequential.

The major problem with this approach resides in estimating a similar wage equation for workers in the informal sector, $(d = 0)$. In the simultaneous case, it is not possible to derive a correction based on the Inverse Mills ratio to estimate the wage equation for the subsample $(d = 0)$, since instead of two, we would have three types of informal sector workers, not clearly defined (Tunali, 1986). Nevertheless, such correction is readily available for the bivariate probit with sample selection. The expected wage for workers not in the queue $(I^*_{1i} < 0)$ can be written as:

$$E(W^1_{ii}|I^*_{1i} < 0) = Z'_{ii}\gamma_i + \sigma_{1i}\lambda_{3ii} + \eta^1_{ii} \tag{3.17}$$

where, $\lambda_{3ii} = -\phi(\tilde{X}'_{1i}\beta_1)/[1 - \Phi(\tilde{X}'_{1i}\beta_1)]$ and $\eta^1_{ii} = \epsilon_{ii} - \sigma_{1i}\lambda_{3ii}$. On the other hand, the expected wage for workers who are in the queue, but were not chosen from the queue can be written as:

$$E(W^2_{ii}|I^*_{1i} > 0 \quad \text{and} \quad I^*_{2i} < 0) = Z'_{ii}\gamma_i + \sigma^*_{1i}\lambda_{1ii} + \sigma_{2i}\lambda_{4ii} + \eta^2_{ii} \tag{3.18}$$

where, $\lambda_{4ii} = -\phi(\tilde{X}'_{2i}\beta_2)/[1 - \Phi(\tilde{X}'_{2i}\beta_2)]$ and $\eta^2_{ii} = \epsilon_{ii} - \sigma_{1i}\lambda_{1ii} - \sigma_{2i}\lambda_{4ii}$

As for the Abowd and Farber bivariate probit, Mengistae (1998) shows that the expected wage for workers with $d = 0$ can be estimated as the weighted average between the expected wage for workers with $I^*_{1i} < 0$ (not in the queue workers) and the expected wage for workers with $I^*_{1i} > 0$ and $I^*_{2i} < 0$ (in the queue but not chosen from the queue workers)

The overall expected wage for the informal sector can be written as $(1-\pi)E(W^1_{ii}|I^*_{1i} < 0) + \pi E(W^2_{ii}|I^*_{1i} > 0 \quad \text{and} \quad I^*_{2i} < 0)$[25] or:

$$E(W_{ii}) = Z'_{ii}\gamma_i + \sigma_{1i}\lambda_{3ii} + \delta_1\lambda^*_{1i} + \delta_2\lambda_{4ii} + \eta_{ii} \tag{3.19}$$

where, $\lambda^*_{1i} = \lambda_{1ii} - \lambda_{3ii}$, $\delta_1 = \pi\sigma_{1i}$, $\delta_2 = \pi\sigma_{2i}$ and $\eta_{ii} = (1-\pi)\eta^1_{ii} + \pi\eta^2_{ii}$.

The bivariate probit estimates from the sequential model estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ are used in (3.14) and (3.15) to approximate the unknown Inverse Mills ratio.

The standard errors of the estimated parameters for both formal and informal workers need to be corrected to account for the heterogeneity introduced by the λ_i's and for the fact that they are first-step estimated variables and not an observed variable, we do that following the adaptation of Mengistae (1998) for Ham's (1982)

[25]Where π corresponds to the proportion of informal workers who are in the queue and were not chosen from it and $(1 - \pi)$ is the proportion of informal workers who are not in the queue. Note that π approximates the size of the queue in this approach.

methodology for the correction of the standard errors of estimated parameters from equations with two selectivity criteria.

An interesting feature of this approach for the estimation of the wage equation for the informal sector is that we can test directly the existence of a queue for formal jobs and get the proportion of (informal) workers queueing, π, for a formal job from this equation. Mengistae (1998) points out that a test for the absence of a job queue, $\pi = 0$, can be based on the asymptotic joint significance test of δ_1 and δ_2. Similarly, a test for a universal job queue, $\pi = 1$, can be done by testing the hypothesis that $\sigma^*_{1i} = \delta_1$. The rejection of both hypothesis means that there is a partial queue for formal jobs.

The third step of the endogenous switching regression model consists in estimating the "in the queue" and the "chosen from queue" equations through a structural bivariate probit. The procedure is based on the inclusion of an additional regressor - the wage difference between the expected (offered) wage that the worker would face in the formal sector and his/her expected (offered) wage in the informal sector - in the "in the queue" equation. We can also add the wage that he/she would get in the formal sector conditional on being in the queue in the "chosen from the queue" equation. Note that we also have to exclude the variables that only affect the wages from this equation. The structural bivariate probit can only be meaningfully estimated for the Abowd and Faber approach and for the bivariate probit with sample selection[26] , since there is no estimate for the informal sector wage equation available from the simultaneous model.

A final point to be addressed refers to the choice of appropriate and believable exclusion restrictions. Most papers that deal with formal versus informal sector choice usually assume that variables such as the presence of children or elderly people in the household and other members of the household income play a crucial role in determining the choice of sector, but do not affect the individual's wage. It is argued that the informal sector offers some sort of flexibility in terms of hours and workplace that makes it more desirable for people who have children and cannot pay for childbearing services (Marcoullier et al., 1997).

In order to identify the sector allocation equations in the case of Brazil, we opt to use a wide set of variables that should capture how an individual in a specific household would value a formal job *vis-à-vis* an informal one. In order to do that, we build the following set of variables that are used in the sector allocation equation(s),

[26]In the case of the bivariate probit with sample selection the wage differences between the offered wage in the formal sector and the offered wage in the informal sector are based on different equations for the informal sector, depending on the desire to switch for a formal job.

but not in the wage equations: 1) variables indicating the number of children and elder people per age group in the household; 2) variables indicating number of other household members by specific working status (e.g. number of registered workers, non-registered workers, self-employed, unemployed); 3) last occupational status[27]; 4) income of the other members of the household (per capita).

The assumptions behind these variables are quite straightforward. The use of the number of children as a proxy for home workload tries to capture whether or not the presence of children makes people more willing to get an informal job that could give them more "flexibility" in terms of hours. This is particularly important for women who bear most of the childbearing cost. For men, particularly for the heads of the household, the presence of children could make an unemployment spell more costly, making them more likely to accept an informal job. The variables related to the distribution of occupational status among other household members intend to capture whether or not the presence of formal workers in the household makes other members of the household less demanding in terms of job characteristics and then less resistant to get an informal job. The presence of formal workers in the household would act as an insurance for the other members of the household. But it also could be the case, that once knowing the benefits related to the possession of a registered work-card, other members would become more reluctant in getting a job without registration. The variables related to the recent employment history of the individual, as measured by her last occupation before moving to the current job, try to capture some inertia factors that would make a worker more or less likely to get a formal job. A former holder of a formal job must have less incentive to get an informal job than a former informal worker with similar characteristics. Similarly, employers tend to look at the employment history of the job applicant when making decision whether to hire him/her. Finally, we also incorporate the per capita income of other members of the household in order to assess how others "disposable incomes" affect the willingness to join a formal job.

The variables listed above enter the univariate probit together with human capital variables and industry and regional characteristics. However, in order to identify separately the "in the queue" equation and the "chosen from the queue" equation in a bivariate context, we exclude the following variables from the "chosen from the queue" equation: tenure, variables related to the number of children in the household and older people, occupational status of the other members of the household,

[27]Last occupational status is only available for workers who had switched to the current job less than 5 yeas ago. For this reason we include "being in the current job for more than 5 years" as a variable in this category.

and per capita income of other members of the household. Tenure, or seniority on the current job, has been traditionally included only in the "in the queue" equation[28]. In doing so, we assume that the worker takes into account the loss of specific human capital and of the benefits that accrue from seniority when joining the queue, but that this specific human capital is not useful for the decision of the employer. Experience as a proxy for general human capital should be the variable that employers rely on when taking the hiring decision[29]. We assume that the variables related to the household situation do not affect the decision of the employers with the exception of the marital status, that traditionally has been used as a signal for commitment to work in the literature of economics of personnel. Thus, the number of children, the number of elder people in the household, the working status of other members and the per capita income of the others members of the household only affect the decision to join the queue. Additionally, in order to identify the wage equation we assume that industry affiliation only affects wages, therefore, we exclude the variables related to industry affiliation from the structural bivariate probit.

3.4 Data

The data used in this paper comes from the 1990 Brazilian Annual Household Survey (PNAD)[30]. We chose this specific year because its survey questionnaire had a special supplement where non-registered workers and self-employed workers were asked whether or not they would like to switch to a formal sector job. This question allows us to relax the severe partial observability of both Abowd and Farber (1982) and Poirier (1980) bivariate probit models. Moreover, this supplement also investigate the past employment history of the individuals yielding a much richer data set than the ones usually available.

In order to properly build variables related to the family composition (e.g. the number of other members of the household who are unemployed and the number of children per age group), we exclude from our sample all persons who are not related to the head of the household where they live. This procedure excludes, for instance, domestic servants who live in the house of their employers.

The sample used in the estimation is also restricted to workers who worked more than 20 hours per week and who were between 15 and 64 years old. We dropped

[28]See Abowd and Farber(1982) and Faber(1983).

[29]We are aware that this assumption is a bit controversial and for this reason, we also check the robustness of the results to the inclusion of the tenure in the "chosen from the queue" equation.

[30]See Chapter 2 subsection 2.2.1 for more details on this data set.

from the sample all individuals for whom there were missing observations for any of the variables used in either the bivariate probit or the wage equations. The resultant sample has 60,138 observations from which 43,322 are formal (registered) workers and 16,816 are informal (non-registered) workers. Table 3.1 depicts the means and the standard deviation of all variables used in the analysis for the full sample and for formal workers and informal workers sub-samples.

The log hourly wage wages for informal workers is substantially lower than for formal workers in 1990. Informal workers have lower schooling, only 10% of them have completed the secondary education or more, whereas 33% of formal workers have at least the secondary education. Male, married and white workers are more likely to be formal as well as workers who live in metropolitan areas and in the South or in the Southeast regions. Formal workers have more than the double of the tenure of their informal counterparts (5.3 *versus* 2.3 years). Formal workers on average have more experience[31] than informal workers.

Formal workers are mainly concentrated on the manufacturing sector (31%) and productive services (25%) (banking, telecommunications, transport), whereas informal workers are prevalent in personal, food and lodging services (37%). As for their last occupation within a 5 year-period, formal and informal workers present different patterns. Both groups show a high degree of inertia, 33% of informal workers used to be informal workers in their last occupation and 34% in the case of formal workers were also formal worker on their last job. However, whereas 31% of informal workers are "new entrants", i.e, they had no previous experience in the labour market, only 21% of the formal workers fall in this category. For formal workers 26% are in the current job for more than 5 years, in contrast, only 12% of informal workers are in the current job for more than 5 years[32]. Only 10% of formal workers were informal in the previous job, whereas 18% of the informal workers were formal. Finally, only 3% of formal workers and 5% of informal workers were self-employed before switching to their current job. These figures suggest that informal jobs seem to be the most common entry into the labour market for many young workers and also suggest that the last occupational status is a good predictor of the current one.

The number of children per age group in the household is higher for informal workers than for registered workers for all age groups. The number of people with more than 70 years is also higher for informal workers. The number of other members of the household who are formal workers is higher for registered workers than for

[31] Experience is calculated as $exp = age - 6 - yearsofschooling$.

[32] This figure does not come as a surprise given the large difference in tenure between the two groups.

	Full Sample		Formal Sector		Informal Sector	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Log hourly wage	0.43	*0.999*	0.68	*0.921*	-0.20	*0.914*
Illiterate (base)	0.08	*0.275*	0.06	*0.236*	0.14	*0.349*
Some elementary	0.14	*0.352*	0.12	*0.322*	0.22	*0.411*
Complete elementary	0.36	*0.480*	0.33	*0.472*	0.42	*0.494*
Complete primary	0.16	*0.368*	0.18	*0.381*	0.12	*0.330*
Complete secondary	0.19	*0.391*	0.23	*0.421*	0.08	*0.266*
Complete college	0.06	*0.244*	0.08	*0.274*	0.02	*0.132*
Exp/10	1.86	*1.259*	1.92	*1.234*	1.70	*1.306*
Exp/100	5.03	*6.272*	5.20	*6.151*	4.60	*6.555*
Tenure	4.42	*5.752*	5.25	*6.068*	2.29	*4.140*
Sex (male=1)	0.63	*0.482*	0.65	*0.477*	0.59	*0.492*
Metropolitan	0.51	*0.500*	0.56	*0.497*	0.38	*0.484*
Race (white=1)	0.53	*0.499*	0.57	*0.495*	0.42	*0.494*
New entrants (base)	0.24	*0.428*	0.21	*0.410*	0.31	*0.463*
More than 5 years	0.26	*0.437*	0.31	*0.462*	0.12	*0.329*
Former formal worker	0.30	*0.457*	0.34	*0.475*	0.18	*0.383*
Former informal worker	0.16	*0.368*	0.10	*0.294*	0.33	*0.471*
Former public servant	0.00	*0.069*	0.01	*0.073*	0.00	*0.061*
Former self-employed	0.04	*0.192*	0.03	*0.177*	0.05	*0.225*
#Children 0-3 years	0.96	*1.135*	0.90	*1.066*	1.11	*1.285*
#Children 4-6 years	0.28	*0.542*	0.27	*0.525*	0.32	*0.582*
#Children7-10 years	0.43	*0.691*	0.39	*0.658*	0.52	*0.763*
#Children 11-13 year	0.33	*0.592*	0.29	*0.562*	0.43	*0.654*
# Elder	0.07	*0.272*	0.06	*0.266*	0.07	*0.286*
# Formal	0.70	*0.960*	0.76	*0.996*	0.56	*0.845*
#Self-employed	0.25	*0.522*	0.21	*0.480*	0.34	*0.607*
#Informal	0.38	*0.777*	0.26	*0.604*	0.71	*1.036*
#Unemployed	0.11	*0.372*	0.11	*0.365*	0.12	*0.388*
Per capita income (others)	0.15	*0.287*	0.16	*0.298*	0.11	*0.255*
Married	0.52	*0.500*	0.57	*0.495*	0.38	*0.485*
Northeast (base)	0.24	*0.425*	0.21	*0.405*	0.31	*0.463*
North	0.09	*0.285*	0.08	*0.275*	0.11	*0.308*
Southeast	0.39	*0.487*	0.42	*0.493*	0.31	*0.462*
South	0.17	*0.373*	0.19	*0.392*	0.11	*0.313*
Mid-West	0.12	*0.326*	0.10	*0.306*	0.16	*0.370*
Manufacturing (base)	0.27	*0.444*	0.31	*0.462*	0.17	*0.377*
Constructing	0.08	*0.277*	0.06	*0.246*	0.13	*0.340*
Retail	0.14	*0.350*	0.15	*0.354*	0.13	*0.339*
Lodging, food and personal services	0.18	*0.386*	0.11	*0.310*	0.37	*0.484*
Productive Sector	0.21	*0.409*	0.25	*0.432*	0.12	*0.326*
Social Services	0.11	*0.311*	0.12	*0.329*	0.07	*0.254*
N	60138		43322		16816	

Table 3.1: Descriptive Statistics for 1990: Full Sample, Registered and Non-registered Workers)

informal workers. Similarly, informal workers are much more likely to live with other informal workers and with unemployed people. Formal workers also live in household where the per capita income of other members is higher in comparison to informal workers.

3.5 Results

3.5.1 Univariate Probit Reduced Form

We will use the results provided by the standard endogenous switching regression model with just one selection equation as a benchmark to evaluate the results for the different specifications of bivariate probit discussed in Section 3.3.

The reduced-form probit of the first step in the estimation of the selectivity-corrected wage equations yields a good description of how the variables included in the (reduced form) model affect the sector choice made by workers in the absence of queue for formal jobs. The results on Table 3.2[33] show that all education groups are more likely to be formal workers than the illiterate group (base category). In fact, the probability of choosing the formal sector increases almost monotonically with the educational level. It also increases with experience but with a decreasing rate. Tenure (or seniority in the current job) has a positive impact on the probability of being a formal worker. The coefficients on gender and on race, despite their positive sign, are not statistically significant, hence, it suggests that there is no "discrimination" on the sector choice (allocation) process[34]. Workers who live in metropolitan areas are more likely to choose formal jobs. Workers who were formal workers in the previous job are more likely to be formal in the current one than "new entrants" (base category) as well as workers who have been in the job for more than 5 years and former public servants and self-employed. However, former informal workers have a lower probability than "new entrants" to be formal in the current job. As for the effect of industry affiliation, all sectors seem to have a lower probability of having formal workers than the manufacturing sector (base category), particularly, lodging, food and personal services and the constructing industry.

Workers who live in the Northeast (base category) are less likely to be formal than workers who live in any other region. Married workers are more likely to be

[33] Notice that we report both the coefficient and the marginal effects.

[34] Carneiro and Henley (2001) and Tanuri-Pianto and Pianto (2002) found significant negative effects for both female and black workers. A possible explanations for differences on these results is the fact that our sample only consider full time workers, i.e, workers who work at least 20 hours per week.

	2-STEP				MLE	
dep. var: formal=1	Coeff.	s.e.	Marg. Effect	s.e.	Coeff.	s.e.
Constant	-0.50	*0.039*	-0.15	*0.012*	-0.44	*0.037*
Some elementary	0.08	*0.026*	0.02	*0.007*	0.08	*0.026*
Complete elementary	0.26	*0.025*	0.07	*0.007*	0.24	*0.025*
Complete primary	0.55	*0.030*	0.14	*0.006*	0.51	*0.029*
Complete secondary	0.80	*0.032*	0.19	*0.006*	0.74	*0.031*
Complete college	0.81	*0.044*	0.18	*0.006*	0.68	*0.041*
Exp/10	0.36	*0.021*	0.11	*0.006*	0.34	*0.020*
$Exp^2/100$	-0.06	*0.004*	-0.02	*0.001*	-0.06	*0.004*
Tenure	0.05	*0.002*	0.01	*0.001*	0.05	*0.001*
Sex (male=1)	0.00	*0.015*	0.00	*0.004*	0.00	*0.014*
Metropolitan	0.19	*0.013*	0.06	*0.004*	0.17	*0.013*
Race (white=1)	0.00	*0.014*	0.00	*0.004*	-0.01	*0.014*
More than 5 years	0.32	*0.021*	0.09	*0.006*	0.37	*0.020*
Former formal worker	0.59	*0.019*	0.16	*0.005*	0.62	*0.019*
Former informal worker	-0.16	*0.020*	-0.05	*0.006*	-0.13	*0.019*
Former public servant	0.26	*0.088*	0.07	*0.021*	0.29	*0.077*
Former self-employed	0.07	*0.033*	0.02	*0.009*	0.14	*0.030*
#Children 0-3 years	-0.03	*0.009*	-0.01	*0.003*	-0.02	*0.009*
#Children 4-6 years	0.01	*0.016*	0.00	*0.005*	0.01	*0.015*
#Children7-10 years	-0.02	*0.012*	-0.01	*0.003*	-0.02	*0.011*
#Children 11-13 year	-0.02	*0.011*	-0.01	*0.003*	-0.03	*0.011*
# Elder	-0.05	*0.023*	-0.02	*0.007*	-0.10	*0.023*
# Formal	0.18	*0.007*	0.05	*0.002*	0.15	*0.007*
#Self-employed	-0.07	*0.012*	-0.02	*0.004*	-0.10	*0.011*
#Informal	-0.19	*0.008*	-0.06	*0.002*	-0.21	*0.008*
#Unemployed	0.02	*0.017*	0.01	*0.005*	0.01	*0.015*
Per capita income (others)	-0.16	*0.024*	-0.05	*0.007*	0.15	*0.016*
Married	0.13	*0.016*	0.04	*0.005*	0.11	*0.016*
North	0.08	*0.024*	0.02	*0.007*	0.08	*0.024*
Southeast	0.28	*0.017*	0.08	*0.005*	0.25	*0.017*
South	0.36	*0.022*	0.10	*0.005*	0.32	*0.022*
Mid-West	0.05	*0.022*	0.01	*0.006*	0.01	*0.022*
Constructing	-0.60	*0.024*	-0.21	*0.009*	-0.57	*0.024*
retail	-0.24	*0.021*	-0.07	*0.007*	-0.21	*0.020*
Lodging, food and personal serv.	-0.90	*0.019*	-0.31	*0.007*	-0.86	*0.019*
Productive Sector	-0.07	*0.020*	-0.02	*0.006*	-0.07	*0.020*
Social Services	-0.15	*0.025*	-0.05	*0.008*	-0.14	*0.025*
Log L	-26054.8					
N	60138				60138	

Base categories: Education: Illiterate, Past experience: Entrants; Region: Northeast; Industry: Manufacturing

Table 3.2: Univariate Probit Specification

formal than single workers and the number of children and the presence of older people in the household seem to have a negative impact on the probability of being formal. The "children effect", however, is not significant for the age groups 4 to 6 and 10 to 13 years of age. The presence of formal workers in the household increases the probability of the individual being in the formal sector whereas the presence of informal workers and self-employed reduces it. The number of unemployed workers does not have a significant effect, but it has a positive sign. As for the other member of the household income (per capita) the probit equation suggests a negative impact on the probability of being formal.

3.5.2 Bivariate Probit Reduced Form

Abowd-Farber Bivariate Probit

The reduced form of the Abowd-Farber bivariate probit with partial observability and sequential decision reveals interesting different patterns for the "in the queue" equation and for the "chosen from the queue" equation (Table 3.3, Part A). The probability of being in the queue is not significantly different among different education levels, whereas the probability of being chosen from the queue is significantly higher for workers with a higher educational level, particularly for the groups with complete secondary school or more. The results of the univariate specification discussed in the last subsection disguise the different impact of schooling on the allocation process observed in the bivariate context.

	Abowd-Farber (A)				Poirier (B)				BP with sample selection (C)			
	IQ		CFQ		IQ		CFQ		IQ		CFQ	
dep. var: formal=1	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Constant	0.32	*0.080*	-0.44	*0.064*	0.71	*0.082*	-0.42	*0.062*	1.44	*0.046*	-0.67	*0.037*
Some elementary	-0.06	*0.054*	0.19	*0.044*	-0.10	*0.055*	0.19	*0.043*	-0.01	*0.031*	0.17	*0.026*
Complete elementary	-0.01	*0.052*	0.46	*0.042*	-0.10	*0.053*	0.44	*0.040*	-0.05	*0.030*	0.45	*0.025*
Complete primary	0.02	*0.060*	0.93	*0.050*	-0.13	*0.062*	0.89	*0.048*	0.00	*0.036*	0.86	*0.030*
Complete secondary	0.08	*0.063*	1.30	*0.052*	-0.12	*0.065*	1.28	*0.051*	0.06	*0.038*	1.26	*0.033*
Complete college	0.06	*0.077*	1.26	*0.068*	-0.10	*0.079*	1.26	*0.066*	0.05	*0.050*	1.51	*0.052*
exp/10	0.09	*0.037*	0.49	*0.035*	0.03	*0.038*	0.47	*0.034*	0.08	*0.024*	0.46	*0.021*
$Exp^2/100$	-0.04	*0.007*	-0.06	*0.006*	-0.03	*0.007*	-0.06	*0.006*	-0.04	*0.004*	-0.05	*0.004*
tenure	0.40	*0.012*			0.39	*0.011*			0.03	*0.002*		
Sex (male=1)	0.00	*0.026*	-0.01	*0.027*	0.00	*0.028*	0.03	*0.025*	-0.08	*0.018*	0.12	*0.015*
Metropolitan	0.16	*0.025*	0.25	*0.024*	0.11	*0.026*	0.24	*0.022*	0.08	*0.017*	0.23	*0.014*
Race (white=1)	-0.02	*0.026*	0.03	*0.025*	-0.02	*0.027*	0.03	*0.024*	-0.08	*0.018*	0.09	*0.014*
More than 5 years	-0.45	*0.057*	0.28	*0.030*	-0.46	*0.055*	0.27	*0.029*	0.20	*0.025*	0.41	*0.024*
Former formal worker	0.27	*0.044*	1.02	*0.048*	0.16	*0.044*	0.89	*0.041*	0.51	*0.025*	0.30	*0.020*
Former informal worker	-0.41	*0.050*	0.33	*0.057*	-0.45	*0.053*	0.36	*0.052*	0.21	*0.024*	-0.36	*0.019*
Former public servant	-0.12	*0.123*	1.10	*0.415*	-0.24	*0.129*	0.99	*0.314*	0.13	*0.111*	0.19	*0.097*
Former self-employed	-0.31	*0.060*	0.87	*0.126*	-0.42	*0.063*	0.78	*0.092*	0.13	*0.039*	-0.14	*0.032*
#Children 0-3 years	-0.02	*0.015*			-0.02	*0.014*			-0.02	*0.011*		
#Children 4-6 years	-0.01	*0.026*			-0.02	*0.023*			0.03	*0.019*		
#Children7-10 years	-0.04	*0.019*			-0.04	*0.017*			-0.01	*0.014*		
#Children 11-13 year	-0.01	*0.018*			-0.01	*0.016*			0.02	*0.013*		
# Elder	-0.09	*0.039*			-0.09	*0.036*			0.00	*0.030*		
# Formal	0.25	*0.014*			0.24	*0.013*			0.16	*0.009*		
#Self-employed	-0.08	*0.019*			-0.08	*0.017*			-0.03	*0.014*		
#Informal	-0.23	*0.012*			-0.20	*0.011*			-0.09	*0.009*		
#Unemployed	0.03	*0.026*			0.03	*0.024*			0.08	*0.022*		
Per capita income (others)	-0.21	*0.030*			-0.19	*0.028*			-0.27	*0.020*		
Married	0.04	*0.028*	0.14	*0.027*	0.02	*0.029*	0.13	*0.025*	0.03	*0.019*	0.12	*0.016*
North	0.07	*0.047*	0.07	*0.041*	0.05	*0.048*	0.08	*0.040*	-0.22	*0.030*	0.20	*0.025*
Southeast	0.19	*0.033*	0.34	*0.031*	0.13	*0.034*	0.31	*0.029*	-0.09	*0.023*	0.41	*0.018*
South	0.32	*0.041*	0.37	*0.038*	0.26	*0.043*	0.34	*0.036*	-0.07	*0.029*	0.52	*0.024*
Mid-West	0.03	*0.041*	0.08	*0.040*	0.02	*0.043*	0.10	*0.038*	-0.21	*0.026*	0.22	*0.022*
Constructing	0.24	*0.068*	-1.11	*0.044*	0.49	*0.070*	-1.08	*0.043*	-0.13	*0.032*	-0.67	*0.024*
Retail	-0.08	*0.036*	-0.37	*0.038*	-0.03	*0.038*	-0.33	*0.037*	-0.21	*0.027*	-0.18	*0.022*
Lodging, food and services	-0.48	*0.040*	-1.10	*0.036*	-0.18	*0.046*	-1.03	*0.035*	-0.44	*0.024*	-0.75	*0.020*
Productive Sector	-0.13	*0.032*	-0.07	*0.037*	-0.12	*0.034*	-0.05	*0.036*	0.00	*0.026*	-0.10	*0.022*
Social Services	-0.27	*0.040*	-0.07	*0.045*	-0.26	*0.042*	-0.05	*0.044*	-0.09	*0.032*	-0.14	*0.028*
Rho(1,2)					-0.80	*0.029*			-0.96	*0.011*		
Log L	-25059				-24979				-35326			
N	60138				60138				60138			

Table 3.3: Bivariate Probit Specifications

Former formal sector workers have the highest probability of being in the queue, followed by the base category "new entrants"[35]. In contrast, all others "last occupation status" have a higher probability of being chosen from the queue than the "new entrants". These probabilities are particularly high for former public servants and former formal workers. These results suggest that employers prefer workers with any type of experience rather than the inexperienced "new entrants".

The estimates for the effect of gender and race are not significant either in the in the queue equation or in the chosen from the queue equation[36]. Married workers are not more likely to be in the queue than single ones. However, they are more likely to be chosen from it. Workers who live in metropolitan areas are more likely to both be in the queue and be chosen from the queue. Workers who live in the Northeast (base category) seems to be both less likely to be in the queue and less likely to be chosen from the queue. As for industry affiliation, manufacturing workers (base category) are more likely to be in the queue in relation to all sectors except for the constructing sector, and also more likely to be chosen from the queue than the other sectors[37].

As for the variables that are only included in the "in the queue" equation the results are not much different from the ones of the univariate specification. Tenure increases the probability of being in the queue for formal jobs. The variables related to number of children per age group, if anything, tend to reduce the probability of being in the queue[38]. The presence of older people has a statistically significant negative effect. The number of other members in the household who are formal workers increases the probability of being in the queue, whereas the number of other informal members decreases it. The presence of self-employed also decreases that probability, whereas the number of unemployed has a positive but not significant effect. Finally, per capita income of other members of the household has a negative effect on the probability of being in the queue.

[35]Surprisingly, former informal workers have the lowest probability of being in the queue.

[36]It is worth noting that the sign of the estimates indicate a positive effect of being male and white on the probability of being chosen from the queue, and a negative effect on the the probability of being in the queue

[37]Actually the coefficient for the productive sector and for the social services are not statistically significant at 5% despite being negative.

[38]The only statistically significant coefficient in this group of variables is the one for the number of children between 10 and 12 years old, for all other categories the sign is negative, but never significant.

Poirier Bivariate Probit

Table 3.3 - Part B, displays the results for Poirier's bivariate probit with partial observability and simultaneous decision. As most of the estimates are quite similar to the sequential decision specification in terms of sign and magnitude, we will only comment on the differences between the two models.

Some intermediate groups of education show a lower probability of being in the queue than the illiterate group (base category), whereas the probability of being chosen from the queue continues to grow almost monotonically with education. Male workers are neither more likely to be chosen from the queue nor more likely than female workers to queue for formal jobs. However, the sign of the coefficient for male workers in the "chosen from the queue" is positive, whereas it is negative in the "in the queue" equation. Similarly, the coefficients on race (white=1) continue to be not significant either in the "in the queue" equation (negative) or in the "chosen from the queue" equation (positive).

The correlation between the two equations, "in the queue" and "chosen from the queue", is negative (-0.80) and statistically significant. This means that unobservables that make a worker more likely to be in the queue make him/her less likely to be chosen from the queue and vice-versa.

Bivariate Probit with Sample Selection

The 1990 special supplement of the Brazilian household survey (PNAD) asked self-employed and informal (non-registered) workers whether or not they would like to switch to a formal (registered) job. Whereas roughly 70% of non-registered workers answered that they would like to, only 30% of the self-employed gave a positive answer. This rough statistics suggests that the desire for a formal job is much more common among informal (non-registered) workers than among self-employed. This fact supports the argument we used in the Chapter 1 that our definition of informal sector is more attached to the labour market *strictu sensu* than the definitions that include self-employed, small employers and non-remunerated workers.

Table 3.4 displays some descriptive statistics for the two types of informal sector workers: potential switchers and potential non-switchers. Among the interesting differences between these two groups we observe that those who want to switch to the formal sector earn less than the ones who do not want to. They are also less educated, less experienced and have a lower amount of seniority. They are predominantly non-white and were also informal workers in their previous job. They have more children and live with other informal workers. The per capita income of the other members of the household is lower for them than for non-switchers.

The 1990 supplement also asked non-switchers about the reasons for staying in the informal sector. Only 9% argued that they would earn more as informal workers. Another 9% argued that they prefer an informal job because of either domestic work or time flexibility. The great majority 65% argued that they simply were satisfied with their current job. This response may in reality represent a mix of satisfaction with their wage and/other advantages offered by informal jobs such as time flexibility. A tiny proportion of 7% revealed that they would not like to move to a registered job because they do not have the necessary requirements to get a job in that sector. This self-selection answer suggests that this group of workers lost their interest in joining the queue after spending some time on it.

In order to assess our previous results of the reduced-form bivariate probit with partial observability, we estimate a bivariate probit with sample selection exploiting the additional information provided by the "would you like to switch to a formal job?" question. We assume that workers who answered positively this question and workers who are in the formal sector are in the queue. Workers who answered the question negatively are assumed not to be in the queue. In doing that we get rid of the severe partial observability and end up with a censored dependent variable.

The results of the bivariate probit with sample selection reported in Table 3.3 - Part C[39] are not very different from the previous ones. This is reassuring because these specifications have different dependent variables. In fact, the most remarkable differences refer to the significance of the estimated parameters for race and gender in both equations, and to the sign of the variables related to "last occupation status" in both equations "in the queue" and "chosen from the queue".

Female and non-white workers are more likely to be in the queue, but are less likely to be chosen from the queue. These results, differently from the previous specifications, are statistically significant and point to the existence of gender and race "discrimination" in hiring policies in the formal sector. Workers in the Northeast now are more likely to be in the queue than workers from any other region, but less likely to be chosen from the queue.

"New entrants" are the least likely category to be in the queue for formal jobs. However, they are not less likely to be chosen from the queue than informal workers. The latter seems to display some characteristics that make them less attractive to formal employers, even when compared to inexperienced workers. These results

[39]Note that the dependent variable for the "in the queue" and the "chosen from the queue" equation are different for the bivariate probit with sample selection are different, unlike the case for the Abowd and Farber model and the Poirier model. This is due to the fact that we observe who is the queue and who is not.

	Potential Switchers		Potential Non-switchers	
	Mean	Std.Dev.	Mean	Std.Dev.
log hourly wage	-0.37	0.81	0.28	1.02
illiterate (base)	0.15	0.35	0.12	0.33
some elementary	0.23	0.42	0.18	0.38
complete elementary	0.44	0.50	0.39	0.49
complete primary	0.12	0.32	0.14	0.35
complete secondary	0.06	0.24	0.13	0.33
complete college	0.01	0.09	0.04	0.21
exp/10	1.57	1.21	2.06	1.48
$Exp^2/100$	3.93	5.79	6.45	8.01
tenure	1.78	3.31	3.68	5.62
sex (male=1)	0.58	0.49	0.62	0.48
metropolitan	0.36	0.48	0.42	0.49
race (white=1)	0.39	0.49	0.53	0.50
new entrants (base)	0.31	0.46	0.31	0.46
more than 5 years	0.09	0.29	0.22	0.41
former formal worker	0.18	0.38	0.18	0.39
former informal worker	0.37	0.48	0.22	0.42
former public servant	0.00	0.05	0.01	0.08
former self-employed	0.05	0.22	0.06	0.24
#children 0-3 years	1.17	1.32	0.95	1.16
#children 4-6 years	0.34	0.60	0.27	0.53
#children7-10 years	0.55	0.78	0.44	0.71
#children 11-13 year	0.46	0.67	0.33	0.59
# elder	0.08	0.29	0.07	0.26
# formal	0.57	0.86	0.53	0.81
#self-employed	0.36	0.62	0.28	0.56
#informal	0.77	1.08	0.53	0.86
#unemployed	0.14	0.40	0.09	0.34
per capita income (others)	0.09	0.14	0.18	0.43
married	0.34	0.47	0.49	0.50
Northeast (base)	0.35	0.48	0.21	0.40
North	0.10	0.30	0.12	0.32
Southeast	0.29	0.46	0.35	0.48
South	0.10	0.30	0.15	0.35
Mid-West	0.16	0.37	0.18	0.38
Manufacturing (base)	0.17	0.37	0.18	0.39
constructing	0.15	0.36	0.09	0.29
retail	0.12	0.33	0.15	0.36
lodging, food and personal services	0.39	0.49	0.33	0.47
Productive Sector	0.11	0.31	0.15	0.36
Social Services	0.06	0.24	0.09	0.29
N	12364		4478	

Table 3.4: Descriptive Statistics for Informal Sector Workers: 1990

suggest that accepting an informal job may worsen the individual chances to get a formal job in the next period conditional on all the variables included in the model. Another different result is the fact that this group is no longer the least likely to be "in the queue", according to this specification, they are only less likely than former formal workers to be "in the queue".

Similarly to the bivariate probit with simultaneous decision, the correlation between the selection equation ("in the queue equation") and the chosen from the queue equation is negative, -0.96, meaning that unobservables that make workers more likely to be in the queue make them less likely to be chosen from the queue.

The last two models show that we cannot accept the hypothesis that both equations, in the queue and chosen from the queue, are not correlated. Actually they are strongly correlated. This means that one of the assumptions of the Abowd and Farber (1982) model does not hold. The choice between Poirier model and the bivariate probit with sample selection depends on how one prefers to model the selection process either as a simultaneous or as sequential process.[40]

3.5.3 Job Queue Tests and the Length of the Queue

Applying the Abowd and Farber (1982) likelihood ratio test (LR) to the hypothesis of a universal queue and to the hypothesis of no queue, we are able to reject both models in favour of a complete bivariate specification in the three specifications that we adopt. As the univariate probit model is not nested in the bivariate specification, we follow the Abowd and Farber (1982) procedure and test those hypotheses by imposing restrictions on the bivariate specifications. For the non-queue model, we impose the restriction that all variables in the "chosen from the queue" equation are zero with the exception of a positive constant that is sufficiently high to make all workers eligible for a registered job. Similarly, when testing the universal queue hypothesis, we impose the restriction that all variables in the "in the queue" equation are zero with the exception of a positive constant. Table 3.5 shows that for all specifications we are able to reject both the universal queue hypothesis and the no queue hypothesis. These results suggest that the bivariate specification is the most appropriate procedure to describe the sector allocation process.

Table 3.6 reports the average probability of being in the formal sector (PF), of being in the queue for a formal job (PIQ) and of being chosen from the queue conditional on being in the queue (PCF) that we get from the three bivariate models

[40] Despite rejecting Abowd and Farber model we will still report the results as this is the standard procedure found in the job queue literature.

described above for some selected characteristics. It also contains an estimate of the length of the queue (q) and the probability of being a formal worker given by the univariate probit. The first interesting thing to note is that the probability of being chosen from the queue, in general, is lower in the bivariate probit with sample selection (censored probit) than in both sequential and simultaneous bivariate probit with partial observability. The only exceptions are workers with at least college education and workers who are in the current job for more than 5 years. For these workers the probability of being chosen from the queue in the specification of the bivariate probit with sample selection is higher than in the other partial observability specifications. Another relevant difference among the results yielded by different specifications is the extremely low probability of being chosen from the queue (conditional on being in the queue) for former informal workers in the bivariate probit with sample selection. This category presents the highest estimate of the length of the queue, 2.11, meaning that for each worker in the formal sector with such characteristics there are 2.11 workers willing to get a formal job[41].

Overall, the different specifications yield an estimate of the length of the queue that varies from 1.20 (Abowd-Farber bivariate probit) to 1.30 (bivariate probit with sample selection). These estimates mean that for each worker in the formal sector there are something around 1.20 or 1.30 workers queuing for formal jobs. However, the interesting point of this analysis is to evaluate how the probability of being chosen from the queue (conditional on being in the queue) varies according to different characteristics. In this regard, female, non-white, former informal workers, "new entrants" and workers with low schooling are the groups who once "in the queue" have the lowest probability of being chosen from it. As pointed out above, being a former informal worker seems to be the most damaging handicap that one can have in the labour market. A spell in an informal job acts as a "scar" for the workers who experience it, signalling some characteristics that are not valued by formal employers. The fact that the probability of being chosen from the queue is even lower than for "new entrants" is quite revealing.

[41]Note that this result should be expected given the different signs of the variable "last occupation as informal worker" in the "in the queue" equation and in the "chosen from the queue" equation in the bivariate probit with sample selection when compared to the other specifications.

	Abowd-Farber		Poirier		Sample selection	
	Universal queue	No queue	Universal queue	No queue	Universal queue	No queue
LR	5081.34	3510.06	5243.27	3672	19470.04	35778.08
p-value	0	0	0	0	0	0
critical value	50.6	37.65	50.6	37.65	50.6	37.65

Table 3.5: LR Test for Universal Queue and No Queue Hypotheses

3.5.4 Wage Equations for Formal and Informal Workers

Univariate-based corrections

The wage equations corrected for selectivity using the univariate reduced-form probit either by the two-step OLS procedure or by the maximum likelihood procedure

	Abowd-Farber		Poirier		Sample Selection		Univariate	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
All								
PCF	0.84	*0.20*	0.80	*0.24*	0.77	*0.24*		
PIQ	0.84	*0.20*	0.87	*0.16*	0.92	*0.06*		
PF	0.72	*0.26*	0.72	*0.26*	0.72	*0.24*	0.72	*0.25*
Q	1.19		1.25		1.30			
Men								
PCF	0.85	*0.19*	0.82	*0.22*	0.79	*0.22*		
PIQ	0.85	*0.19*	0.88	*0.15*	0.93	*0.05*		
PF	0.74	*0.24*	0.74	*0.24*	0.74	*0.22*	0.74	*0.23*
Q	1.18		1.22		1.26			
Women								
PCF	0.82	*0.21*	0.77	*0.27*	0.73	*0.27*		
PIQ	0.81	*0.23*	0.85	*0.17*	0.92	*0.07*		
PF	0.68	*0.28*	0.68	*0.29*	0.68	*0.27*	0.69	*0.27*
Q	1.23		1.30		1.37			
White								
PCF	0.88	*0.17*	0.85	*0.20*	0.83	*0.20*		
PIQ	0.87	*0.18*	0.89	*0.14*	0.92	*0.06*		
PF	0.78	*0.23*	0.78	*0.23*	0.78	*0.21*	0.78	*0.22*
Q	1.14		1.17		1.20			
Non-white								
PCF	0.79	*0.22*	0.74	*0.26*	0.70	*0.26*		
PIQ	0.80	*0.22*	0.85	*0.17*	0.92	*0.06*		
PF	0.65	*0.27*	0.66	*0.28*	0.65	*0.26*	0.66	*0.26*
Q	1.27		1.34		1.43			
Entrants								
PCF	0.72	*0.24*	0.68	*0.27*	0.69	*0.25*		
PIQ	0.86	*0.18*	0.91	*0.12*	0.90	*0.07*		
PF	0.64	*0.27*	0.64	*0.28*	0.63	*0.24*	0.64	*0.25*
Q	1.40		1.46		1.45			
+ 5 Years								
PCF	0.89	*0.14*	0.89	*0.15*	0.92	*0.10*		
PIQ	0.96	*0.12*	0.97	*0.09*	0.94	*0.05*		
PF	0.86	*0.18*	0.86	*0.17*	0.86	*0.12*	0.86	*0.15*
Q	1.12		1.13		1.09			
Former formal								
PCF	0.95	*0.08*	0.93	*0.10*	0.87	*0.12*		
PIQ	0.88	*0.11*	0.90	*0.09*	0.95	*0.03*		
PF	0.83	*0.14*	0.83	*0.13*	0.83	*0.13*	0.83	*0.14*
Q	1.06		1.08		1.15			
Former informal								
PCF	0.71	*0.21*	0.60	*0.28*	0.47	*0.25*		
PIQ	0.57	*0.22*	0.67	*0.18*	0.90	*0.06*		
PF	0.43	*0.23*	0.43	*0.25*	0.43	*0.24*	0.43	*0.22*
Q	1.42		1.67		2.11			
Illiterate								
PCF	0.68	*0.22*	0.61	*0.26*	0.58	*0.26*		
PIQ	0.74	*0.26*	0.81	*0.20*	0.88	*0.08*		
PF	0.52	*0.26*	0.52	*0.27*	0.52	*0.25*	0.52	*0.25*
Q	1.46		1.63		1.74			
Some elementary								
PCF	0.73	*0.23*	0.67	*0.28*	0.63	*0.28*		
PIQ	0.76	*0.24*	0.83	*0.19*	0.90	*0.06*		
PF	0.58	*0.28*	0.58	*0.29*	0.58	*0.26*	0.58	*0.27*
Q	1.36		1.48		1.58			
								(cont)

Table 3.6: Probabilites of a) Being Chosen From the Queue Conditional on Being in the Queue (PCF), b) Being in the Queue (PIQ), c) Being in the Formal Sector (PF); and Length of the Queue (Q).

	Abowd-Farber		Poirier		Sample Selection		Univariate (cont)	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Complete elementary								
PCF	0.80	*0.20*	0.76	*0.25*	0.72	*0.24*		
PIQ	0.81	*0.21*	0.86	*0.16*	0.92	*0.06*		
PF	0.67	*0.26*	0.67	*0.26*	0.67	*0.24*	0.67	*0.25*
Q	1.25		1.32		1.39			
Complete primary								
PCF	0.90	*0.13*	0.87	*0.15*	0.83	*0.16*		
PIQ	0.87	*0.16*	0.89	*0.13*	0.94	*0.05*		
PF	0.78	*0.20*	0.78	*0.19*	0.78	*0.17*	0.78	*0.19*
Q	1.11		1.14		1.20			
Complete secondary								
PCF	0.96	*0.06*	0.95	*0.07*	0.93	*0.08*		
PIQ	0.92	*0.12*	0.93	*0.11*	0.95	*0.04*		
PF	0.88	*0.14*	0.89	*0.13*	0.88	*0.09*	0.88	*0.12*
Q	1.04		1.05		1.08			
Complete college								
PCF	0.97	*0.04*	0.97	*0.05*	0.97	*0.04*		
PIQ	0.95	*0.10*	0.95	*0.09*	0.95	*0.05*		
PF	0.92	*0.11*	0.92	*0.10*	0.92	*0.07*	0.92	*0.09*
Q	1.03		1.03		1.03			

yield very similar results[42]. Both results shown in Table 3.7, part A and B - are quite robust in showing the existence of selectivity in both the formal sector and the informal sector and in indicating an overestimation of most parameters in the standard OLS procedure (Table 3.7, part C). In the selectivity-corrected OLS procedure, the negative and statistically significant coefficient (-0.20) of the Inverse Mills ratio for the formal sector equation suggests that individuals who are less likely to be registered workers benefit more of this status than the ones who are more likely to be registered. There is a negative correlation (-0.32) between the residual of the probit specification and the residual of the wage equation. This result does not support the hypothesis that individuals allocate themselves into the sectors where they have comparative advantage. Similar results are obtained in the ML approach, since the correlation between the residuals of the selection equation and the wage equation is also negative and statistically significant, -0.53. Tannuri-Pianto and Pianto (2002) find similar results for Brazil using a semiparametric approach. They argue that the negative effect of the selectivity parameters on wages indicates that workers choose to join the formal sector due to non-wage benefits associated with this kind of job that are not captured by the wage equation. The shortcoming of this sort of explanation is that it assumes that every worker who prefers a formal to an informal job is able to get it, i.e, there is no queue for formal jobs. Another important fact that one should not overlook is that workers who are less likely to be formal workers are the ones who tend to earn relatively more once in the formal

[42]The selection equation in the ML estimation also do not differ much from the probit estimation of the two-step procedure (see Table 3.2), only the coefficients on race and on per capita income display a different sign. However, only the latter is statistically significant in both equations.

sector.

As for the wage equation of the three-step OLS procedure for the informal sector the coefficient of the Inverse Mills ratio is negative and significant, -0.21. This result indicates the presence of selectivity in the informal sector. The workers who are less likely to be in the formal sector are the ones who command higher than expected wages in the informal sector. The residuals of the selection equation (probability of choosing the formal sector) and the residuals of the wage equation are negatively correlated, -0.31[43]. One can argue that this result suggests that informal workers choose to join the informal sector due to comparative advantages in that sector (Carneiro and Henley, 2001)[44]. However, in order to conclude that workers select themselves in the sector where they are have comparative advantage, we should also find a similar result for the formal sector equation, which is not the case[45].

The estimates of the effect of human capital variables - groups of education, experience, experience squared, tenure - in the formal sector equation are overestimated in the non-corrected OLS. The ML equation estimates show a somewhat higher bias in the OLS procedure than the two-step procedure. The estimates for the informal sector are also overestimated, but now in some cases, the two-step procedure estimates are even lower than their ML correspondents.

It is worth noting that even the selectivity-corrected estimates show strong returns to education for both formal and informal workers. Thus, it is not correct to treat the 'informal' sector as a secondary sector in which there is no return to education[46]. However, it seems that tenure is much more rewarded in the formal than in the informal sector. The results also show the existence of a wage premium for married workers.

Bivariate-based corrections

The wage equations corrected by the double selectivity criteria, for both formal and informal workers also show that the non-corrected OLS procedure renders overestimated estimates for most coefficients in the case of sequential decision models, but

[43]The ML procedure also gives similar results for the correlation between the residual of the selection equation and the residual of the wage equation for the informal sector, the correlation is negative and statistically significant, -0.33.

[44]Carneiro and Henley (2001) estimate an endogenous switching regression model for formal and informal sector using the Brazilian annual household survey for 1999 and find similar results in terms of sign and significance of the selectivity terms. However, they focus their analysis on the coefficient of the Inverse Mills ratio for the informal sector wage equation, and overlook the implication of the sign of the coefficient of the Inverse Mills ratio for the formal sector.

[45]See Yamada(1996) and Maddala (1983) for a discussion of this point.

[46]Similar results are found for different Latin American countries. See Yamada (1996) and Saavedra and Chong (1999).

	SELECTIVITY CORRECTED - OLS (A)				OLS (C)			
	FORMAL		INFORMAL		FORMAL		INFORMAL	
dep. var.: log hourly wage	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-0.96	*0.025*	-1.80	*0.028*	-1.14	*0.019*	-1.76	*0.027*
Some elementary	0.14	*0.015*	0.15	*0.019*	0.16	*0.015*	0.17	*0.018*
Complete elementary	0.32	*0.015*	0.36	*0.019*	0.36	*0.014*	0.42	*0.018*
Complete primary	0.64	*0.017*	0.64	*0.026*	0.70	*0.016*	0.75	*0.023*
Complete secondary	1.12	*0.018*	1.07	*0.031*	1.20	*0.016*	1.23	*0.027*
Complete college	1.93	*0.020*	1.89	*0.047*	2.01	*0.019*	2.05	*0.044*
Exp/10	0.30	*0.010*	0.36	*0.017*	0.33	*0.010*	0.43	*0.016*
$Exp^2/100$	-0.05	*0.002*	-0.05	*0.003*	-0.06	*0.002*	-0.06	*0.003*
Tenure	0.03	*0.001*	0.01	*0.002*	0.03	*0.001*	0.02	*0.001*
Sex (male=1)	0.30	*0.007*	0.38	*0.012*	0.30	*0.007*	0.39	*0.012*
Metropolitan	0.09	*0.006*	0.19	*0.012*	0.11	*0.006*	0.23	*0.011*
Race (white=1)	0.16	*0.007*	0.11	*0.012*	0.16	*0.007*	0.11	*0.012*
Married	0.17	*0.007*	0.21	*0.014*	0.18	*0.007*	0.23	*0.013*
North	0.42	*0.012*	0.48	*0.019*	0.43	*0.012*	0.49	*0.019*
Southeast	0.28	*0.009*	0.31	*0.015*	0.31	*0.008*	0.36	*0.014*
South	0.27	*0.011*	0.35	*0.021*	0.31	*0.010*	0.42	*0.020*
Mid-West	0.41	*0.012*	0.46	*0.017*	0.41	*0.011*	0.47	*0.016*
Constructing	-0.01	*0.014*	0.14	*0.021*	-0.06	*0.013*	0.06	*0.019*
Retail	-0.21	*0.010*	-0.02	*0.020*	-0.23	*0.010*	-0.05	*0.019*
Lodging, food and services	-0.22	*0.013*	-0.13	*0.020*	-0.31	*0.011*	-0.25	*0.016*
Productive Sector	0.02	*0.008*	0.10	*0.020*	0.02	*0.008*	0.10	*0.020*
Social Services	-0.16	*0.011*	-0.06	*0.025*	-0.18	*0.011*	-0.08	*0.025*
Inverse Mills ratio (lambda)	-0.20	*0.018*	-0.21	*0.022*				
R^2	0.56		0.46		0.55		0.46	

	MLE (B)			
	FORMAL		INFORMAL	
dep. var.: log hourly wage	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-0.83	*0.023*	-1.80	*0.027*
Some elementary	0.13	*0.017*	0.15	*0.019*
Complete elementary	0.30	*0.016*	0.36	*0.019*
Complete primary	0.60	*0.018*	0.63	*0.025*
Complete secondary	1.06	*0.018*	1.05	*0.030*
Complete college	1.88	*0.020*	1.86	*0.040*
Exp/10	0.27	*0.010*	0.36	*0.017*
$Exp^2/100$	-0.05	*0.002*	-0.05	*0.003*
Tenure	0.02	*0.001*	0.01	*0.001*
Sex (male=1)	0.30	*0.007*	0.38	*0.012*
Metropolitan	0.08	*0.006*	0.18	*0.012*
Race (white=1)	0.16	*0.007*	0.11	*0.012*
Married	0.16	*0.007*	0.21	*0.013*
North	0.42	*0.011*	0.47	*0.018*
Southeast	0.26	*0.009*	0.30	*0.015*
South	0.25	*0.011*	0.35	*0.022*
Mid-West	0.41	*0.011*	0.46	*0.017*
Constructing	0.03	*0.014*	0.14	*0.022*
Retail	-0.20	*0.010*	-0.03	*0.019*
Lodging, food and services	-0.16	*0.012*	-0.13	*0.020*
Productive Sector	0.03	*0.008*	0.10	*0.019*
Social Services	-0.15	*0.011*	-0.06	*0.024*
$\rho(1,u)$	-0.53	*0.015*	-0.33	*0.028*

Table 3.7: Wage Equations: Univariate-based corrections

not in the simultaneous decision model. In the case where the correction is based on the Abowd-Farber bivariate probit with partial observability and sequential decision (Table 3.8, part A) all estimates of selectivity correction parameters are negative and significantly different from zero[47]. Therefore, both the workers less likely to be in the queue and the workers less likely to be chosen from the queue are the ones who tend to earn more than would be expected, given their productive characteristics, in the formal sector. Differently, the informal workers who are less likely to be in the queue for formal jobs and those who are less likely to be chosen from the queue are the ones who benefit more from the informal job. As most results for the estimated wage equations remain unchanged we avoid repetition and analyse only the result of the job queue test suggested in Mengistae (1998). The hypothesis of no job queue ($\delta_1 = \delta_2 = 0$) is rejected at a standard 5% level of significance. The calculated F statistics is F(1, 16791)= 77.27 with a p-value of 0.000. On the other hand, we cannot reject the hypothesis of a universal queue for formal jobs $\pi = 1$. The F statistic for the restriction that $\sigma_{1i} = \delta_1$ is 0.009 with a p-value of 0.9754[48].

As for the correction based on the simultaneous model (Poirier bivariate probit) we can only estimate a meaningful equation for the formal sector (see Section 3.3). The estimates of the human capital variables (Table 3.8, part B), if anything, seem to be underestimated in the non-corrected OLS regression. The selectivity coefficient for the "chosen from the queue equation", indicates that workers more likely to be chosen from the queue are the ones who earn more in the formal sector, whereas the selectivity coefficient for the "in the queue" equation indicates that workers more likely to be in the queue earn less than would be expected given their productive characteristics. Note that this result is quite different from the one yielded by the Abowd-Farber bivariate probit with partial observability and sequential decision model. The fact that the difference occurs in the "chosen from the queue" equation is not surprising since this probability is defined over different populations in these two models as shown in Section 3.3.

The results for the Bivariate Probit with sample selection (Table 3.8, part C) indicates that the results from the non-corrected OLS regressions are overestimated. The three different wage regressions derived from this approach reveal a somewhat

[47] $\sigma_{1f} = -0.19$ and $\sigma_{2f} = -0.07$ in the formal sector wage equation and $\sigma_{1i} = -0.18$ and $\sigma_{2i} = -0.18$ in the informal sector wage equation

[48] As Mengistae (1998) points out, his methodology for testing the job queue using the unconditional wage equation for workers who were not chosen from the queue or were not in the queue is completely independent of Abowd and Farber (1982) test for job queue. In our case, it seems that both methods point to the rejection of the "in the queue" univariate specification for the correction of the wage equations. However, they yield different results in relation to the existence of universal queue.

	ABOWD-FABER (A)				POIRIER (B)	
	FORMAL		INFORMAL		FORMAL	
dep. var.: log hourly wage	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-1.02	0.025	-1.69	0.037	-1.14	0.024
Some elementary	0.15	0.015	0.14	0.019	0.16	0.016
Complete elementary	0.34	0.015	0.33	0.020	0.37	0.015
Complete primary	0.67	0.018	0.57	0.029	0.72	0.018
Complete secondary	1.16	0.019	0.97	0.037	1.22	0.019
Complete college	1.97	0.021	1.81	0.051	2.03	0.021
Exp/10	0.31	0.010	0.32	0.019	0.34	0.010
$Exp^2/100$	-0.05	0.002	-0.05	0.003	-0.06	0.002
Tenure	0.02	0.001	0.01	0.004	0.02	0.001
Sex (male=1)	0.30	0.007	0.36	0.012	0.30	0.007
Metropolitan	0.10	0.006	0.16	0.012	0.11	0.006
Race (white=1)	0.16	0.007	0.11	0.012	0.16	0.007
Married	0.17	0.007	0.20	0.014	0.18	0.007
North	0.42	0.012	0.47	0.019	0.43	0.012
Southeast	0.29	0.009	0.28	0.016	0.31	0.009
South	0.28	0.011	0.32	0.021	0.30	0.011
Mid-West	0.41	0.012	0.45	0.017	0.42	0.012
Constructing	-0.04	0.015	0.19	0.026	-0.09	0.016
Retail	-0.22	0.010	0.00	0.020	-0.23	0.010
Lodging, food and services	-0.25	0.014	-0.06	0.023	-0.32	0.014
Productive Sector	0.02	0.008	0.11	0.020	0.02	0.008
Social Services	-0.17	0.011	-0.05	0.024	-0.17	0.011
σ_{1f}	-0.19	0.016			-0.15	0.017
σ_{2f}	-0.07	0.025			0.08	0.023
σ_{1i}			-0.18	0.018		
δ_1			-0.18	0.021		
δ_2			-0.18	0.021		
R^2	0.56		0.47		0.56	
	BP WITH SAMPLE SELECTION (C)					
dep. var.: log hourly wage	FORMAL		INFORMAL (IQ)		INFORMAL (NIQ)	
Constant	-0.99	0.029	-1.62	0.031	-1.62	0.115
Some elementary	0.14	0.016	0.11	0.021	0.21	0.042
Complete elementary	0.32	0.016	0.26	0.023	0.50	0.040
Complete primary	0.64	0.019	0.47	0.033	0.92	0.049
Complete secondary	1.11	0.020	0.81	0.044	1.37	0.052
Complete college	1.91	0.023	1.64	0.076	2.06	0.069
Exp/10	0.31	0.010	0.28	0.022	0.47	0.032
$Exp^2/100$	-0.06	0.002	-0.04	0.004	-0.07	0.006
Tenure	0.03	0.001	0.01	0.002	0.01	0.002
Sex (male=1)	0.29	0.007	0.35	0.014	0.35	0.025
Metropolitan	0.10	0.007	0.19	0.014	0.18	0.025
Race (white=1)	0.15	0.007	0.07	0.013	0.16	0.025
Married	0.17	0.007	0.17	0.015	0.30	0.026
North	0.40	0.013	0.40	0.021	0.53	0.042
Southeast	0.28	0.009	0.30	0.018	0.23	0.033
South	0.27	0.011	0.34	0.025	0.28	0.041
Mid-West	0.38	0.012	0.47	0.019	0.33	0.038
Constructing	-0.02	0.015	0.15	0.025	0.13	0.046
Retail	-0.23	0.010	-0.08	0.021	0.04	0.040
Lodging, food and services	-0.29	0.013	-0.15	0.024	-0.18	0.038
Productive Sector	0.02	0.009	0.06	0.022	0.20	0.039
Social Services	-0.18	0.011	-0.02	0.027	-0.15	0.047
σ_{1f}	0.31	0.037				
σ_{2f}	-0.24	0.030				
σ^*_{1i}			-0.89	0.123		
σ_{1i}					-0.01	0.054
σ_{2i}			-0.17	0.030		
R^2	0.56		0.40		0.48	

Table 3.8: Wage Equations: Bivariate-based Corrections

different pattern than the ones seen so far. In the case of the formal sector, the positive and significant coefficient on the Inverse Mills ratio from the "in the queue equation", σ_{1f}, suggests that workers more likely to be in the queue have higher wages than expected in the formal sector. Differently, the coefficient on the Inverse Mills ratio from the "chosen from the queue equation",σ_{2f}, suggests that workers more likely to be chosen from the queue earn less than expected. The result for the correction based on the "in the queue" equation differs from the one based on the Abowd-Farber bivariate probit and the result for the correction based on the "in the queue" equation differs from the Poirier bivariate probit.

As for the informal workers we have two different equations. One for the workers who are in the queue and were not chosen from it and one for the workers who were not in the queue. The results for the former group suggest that workers more likely to be in the queue earn lower wages in the informal sector[49], whereas workers less likely to be chosen from the queue are the ones who earn more in the informal sector[50]. This result suggests that informal workers who are "in the queue" for formal jobs tend to have lower wages than expected.

As for the equation for the workers who are not in the queue there is no evidence of the presence of selectivity, the sign of the coefficient on the Inverse Mills ratio from the "in the queue equation" is negative, but not statistically significant,$\sigma_{1i} = -0.01$).

In short, as can be seen in Table 3.9[51], different specifications of the sector allocation process imply different results in terms of the effect of selectivity on earnings. Assuming that the correct model is one with sequential decision (either Abowd-Farber bivariate probit or bivariate probit with sample selection) we can conclude that workers more likely to be "chosen from the queue" are the ones who will benefit less (relatively) from a formal job. This may be indicating some cost minimizing behavior of the employers. The results regarding the "in the queue" equation, however, differ according to the sequential specification adopted.

For the formal sector, both sequential specifications indicate a negative effect of the "chosen from the queue" status on the wage equation. The "in the queue" equation, however, displays a negative effect in the Abowd-Farber specification, and a positive one in the bivariate probit with partial observability.

As for the informal sector, the Abowd-Farber specification, indicates that the "in the queue" status has a positive impact on earnings, whereas the bivariate probit with sample selection shows a negative impact on the wage for informal workers who

[49] $\sigma^*_{1i} = -0.89$
[50] $\sigma_{2i} = -0.17$
[51] Table 3.9 highlights how unobservables of the two sector-selection equations affect wages of the formal and informal sector.

	Predicted wages					
	Abowd-Faber		Poirier	Bivariate Probit with Sample Selection		
	Formal	Informal	Formal	Formal	Informal in the queue	Informal not in the queue
More likely to be in the queue	lower	lower	lower	higher	lower	not sig.
More likely to be chosen from the queue	lower	lower	higher	lower	higher	

Table 3.9: Effect of the correlation between Unobservables in the Choice Equations and in the Outcome Equation

are "in the queue" and a positive, but not statistically significant effect for informal workers who are not it the queue[52].

3.5.5 The Role of Wage Differential on Sector Allocation

Univariate Specification

The structural probit is estimated including the wage differential as an additional regressor and excluding the variables that are assumed to determine only the wage equation (industry dummies). The wage differential is computed as the difference between the offered wage in the formal sector minus the offered wage in the informal sector. The offered wage differs from the conditional wage because it does not include the Inverse Mills ratio in the calculation of the predicted (log) wage[53].This is so, because we are interested in the wage that any member of the population of employees would get, regardless if he/she is actually working in the informal or in the formal sector. Table 3.10 shows both the results based on the three-step OLS and the result based on the maximum likelihood procedure for the univariate case (probit). The results of the OLS-based procedure suggest that the wage differential is the most important variable determining the "choice" of formal sector by workers. The coefficient is positive and statistically significant. The marginal effect is also positive and statistically significant 0.55 *(0.030)*. The marginal effect indicates that the wage differential is the most important variable determining sector allocation in

[52] Note that we can only make this distinction between informal workers in the queue and informal workers not in the queue when using the correction based on the bivariate probit with sample selection.

[53] The unconditional mean formal sector wage premium varies a lot according to the specification adopted: 22% in the OLS specification, 73% in the univariate probit-corrected specification, 105% in the Abowd-Farber bivariate probit-corrected specification; and 93% in relation to informal workers who want to switch to formal sector and -0.3% in relation to informal workers who do not want to switch, according to the bivariate probit with sample selection-corrected specifications. Thus it seems that the OLS specification underestimates the mean wage premium for formal sector workers. However, this mean wage premium is different for different type of informal workers as revealed by the bivariate probit with sample selection in the expected direction.

	OLS		MLE	
dep. var: formal=1	Coeff.	Std.Err.	Coeff.	Std.Err.
(W_f - W_i)	1.83	*0.100*	0.12	0.107
Constant	-2.33	*0.083*	-1.07	0.104
Some elementary	0.12	*0.025*	0.11	0.025
Complete elementary	0.38	*0.024*	0.33	0.025
Complete primary	0.71	*0.029*	0.70	0.029
Complete secondary	0.93	*0.031*	1.01	0.031
Complete college	0.97	*0.043*	1.05	0.043
Exp/10	0.45	*0.021*	0.35	0.022
$Exp^2/100$	-0.06	*0.004*	-0.06	0.004
Tenure	0.03	*0.002*	0.05	0.002
Sex (male=1)	0.26	*0.016*	0.13	0.016
Metropolitan	0.33	*0.016*	0.17	0.017
Race (white=1)	-0.07	*0.015*	0.03	0.015
More than 5 years	0.33	*0.021*	0.33	0.021
Former formal worker	0.59	*0.019*	0.59	0.019
Former informal worker	-0.20	*0.019*	-0.20	0.019
Former public servant	0.27	*0.087*	0.27	0.087
Former self-employed	0.06	*0.032*	0.05	0.032
#Children 0-3 years	-0.04	*0.009*	-0.04	0.009
#Children 4-6 years	0.02	*0.016*	0.02	0.016
#Children7-10 years	-0.02	*0.011*	-0.02	0.011
#Children 11-13 year	-0.02	*0.011*	-0.02	0.011
# Elder	-0.04	*0.023*	-0.04	0.023
# Formal	0.18	*0.007*	0.19	0.007
#Self-employed	-0.07	*0.011*	-0.07	0.011
#Informal	-0.19	*0.008*	-0.19	0.008
#Unemployed	0.02	*0.016*	0.02	0.016
Per capita income (others)	-0.13	*0.024*	-0.13	0.024
Married	0.23	*0.016*	0.17	0.017
North	0.18	*0.024*	0.09	0.024
Southeast	0.31	*0.017*	0.27	0.017
South	0.49	*0.023*	0.36	0.024
Mid-West	0.11	*0.022*	0.02	0.022
Log L	-27347.97		-27515.92	
N	60138		60138	

Table 3.10: Structural Univariate Probit

the univariate model[54]. Nevertheless the other variables of the model continue to be significant and the only noticeable change relative to the reduced form specification is that the coefficient on race is negative and significant, whereas in the reduced form it was positive and statistically insignificant [55].

The results based on the ML procedure, however, show a smaller and statistically non-significant positive effect of the wage differential on the probability of choosing the formal sector. The marginal effect is also small and not significant, 0.04(0.032) As for the other variables, only the coefficient on race shows a different sign from the OLS-based structural probit. In this specification, white workers have a higher probability of being found in the formal sector than non-white workers, but neither the coefficient nor the marginal effects are statistically significant.

[54]If instead of using offered wages, we use conditional wages to calculate the wage differential we get very similar results, but the effect of the wage differential becomes even larger. These results are available upon request.

[55]The coefficient on gender remains positive and turns out to be significant, indicating that male workers are more likely to be found in the formal sector.

Bivariate specifications

In the case of the structural bivariate probit we exclude the variables exclusive to the wage equation (i.e. industry dummies) from the "in the queue" equation, and include the "wage in the formal sector conditional on being in the queue", $(W_f|IQ = 1)$, as an additional regressor in the "chosen from the queue" equation. The results based on the Abowd-Farber structural bivariate probit with partial observability in Table 3.11 (panel A) reveal that for the "in the queue" equation the wage differential plays the most important role in determining the "in the queue" status, its coefficient is positive and significant[56]. The variables related to education that were statistically insignificant in the reduced form, turned out to be negative and significant in some cases. The coefficients on experience and experience squared turn out to be statistically insignificant. The other differences refer to the fact that male workers are now more likely to be "in the queue". The other coefficients did not change much in relation to the previous reduced form results.

As for the "chosen from the queue" equation the results change a lot in relation to the reduced form. The coefficient on the "formal sector wage" conditional on being in the queue is positive and significant[57]. However, one should expect that if formal sector employers were cost minimizers, the sign of this coefficient should be negative. The higher the wage of a worker conditional on other productive attributes, the less likely the employer would be to hire him/her. Besides this awkward result, there were major changes in the sign of other variables such as education groups, sex and race, in most cases in unexpected ways[58]. This fact led us to believe that the "formal sector wage" variable is picking up the effect of human capital-related variables, instead of isolating the effect that a cost-minimization decision would have in the hiring process. Then, we re-estimate the bivariate probit excluding the "formal sector wage" from the "chosen from the queue" equation. The results in Table 3.11 (panel B) are now much more similar to the ones we get from the reduced form of the "chosen from the queue" equation. Moreover, the "in the queue" equation does not change much. The coefficient of the wage differential is positive and significant[59], and despite small changes in magnitudes, the sign of all parameters remain the same as in the previous reduced-form specification.

[56] Its marginal effect is also positive and significant, 0.45 *(0.01)*.

[57] Its marginal effect is also positive and significant, 0.423 *(0.001)*.

[58] According to the results of this specification, female and non-white workers would be more likely to be chosen from the queue than male and white workers. Similarly, workers with complete secondary education or more would be less likely to be chosen from the queue.

[59] Its marginal effect now is still positive and significant, 0.30 *(0.016)*.

	ABOWD-FARBER (A)				ABOWD-FARBER (B)			
	IQ		CFQ		IQ		CFQ	
dep. var: formal=1	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
$(W_f - W_i)$	1.67	0.185			2.72	0.110		
$(W_f \mid IQ = 1)$			1.68	0.043				
Constant	-0.95	0.154	1.27	0.067	-1.60	0.079	-0.89	0.069
Some elementary	-0.17	0.090	-0.07	0.034	-0.14	0.046	0.34	0.049
Complete elementary	-0.15	0.087	-0.17	0.035	-0.10	0.047	0.71	0.046
Complete primary	-0.19	0.098	-0.35	0.049	-0.12	0.060	1.22	0.055
Complete secondary	-0.29	0.107	-0.89	0.066	-0.22	0.070	1.65	0.056
Complete college	-0.19	0.137	-2.31	0.103	-0.20	0.080	1.65	0.070
Exp/10	0.06	0.062	-0.30	0.031	0.17	0.031	0.40	0.040
$Exp^2/100$	0.00	0.013	0.05	0.005	-0.03	0.006	-0.05	0.007
Tenure	1.31	0.063			0.27	0.008		
Sex (male=1)	0.37	0.041	-0.60	0.025	0.33	0.022	-0.03	0.028
Metropolitan	0.27	0.040	0.03	0.018	0.36	0.021	0.18	0.027
Race (white=1)	-0.12	0.042	-0.24	0.020	-0.17	0.022	0.07	0.029
More than 5 years	-0.25	0.110	0.29	0.023	-0.40	0.048	0.34	0.033
Former formal worker	0.26	0.065	0.82	0.026	0.31	0.038	1.37	0.123
Former informal worker	-0.41	0.069	0.20	0.031	-0.43	0.045	0.60	0.106
Former public servant	0.57	0.405	0.33	0.105	-0.01	0.123	0.98	0.494
Former self-employed	0.15	0.111	0.26	0.046	-0.15	0.059	0.77	0.176
#Children 0-3 years	-0.01	0.024			-0.03	0.012		
#Children 4-6 years	-0.02	0.041			0.00	0.021		
#Children7-10 years	-0.06	0.029			-0.03	0.016		
#Children 11-13 year	-0.01	0.028			-0.02	0.015		
# Elder	-0.12	0.060			-0.05	0.033		
# Formal	0.27	0.022			0.21	0.011		
#Self-employed	-0.05	0.030			-0.07	0.016		
#Informal	-0.24	0.017			-0.22	0.011		
#Unemployed	0.04	0.041			0.03	0.022		
Per capita income (others)	-0.20	0.045			-0.17	0.029		
Married	0.08	0.045	-0.15	0.021	0.15	0.023	0.18	0.030
North	0.17	0.079	-0.64	0.035	0.15	0.038	0.14	0.046
Southeast	0.03	0.052	-0.16	0.026	0.13	0.028	0.38	0.035
South	0.28	0.065	-0.10	0.031	0.36	0.033	0.43	0.044
Mid-West	0.12	0.069	-0.66	0.033	0.13	0.035	0.01	0.045
Log L	-26018				-26229			
N	60138							

Table 3.11: Abowd-Farber Bivariate Probit Specification

The results of the structural bivariate probit with sample selection[60] (Table 3.12) show a strong and statistically significant effect of the wage differential on the probability of being in the queue, and also a small, but significant effect of the "formal sector wage" on the probability of being chosen from the queue. The marginal effect of wage differential in the "in the queue" equation is positive (0.018) and significant ($z = 2.00$), the marginal effect of the wage in the formal sector in the "chosen from the queue" equation is also positive and significant, but very small (0.003)[61]. In this latter case, the coefficients on the other variables do not seem to be much affected by the inclusion of this variable. Education, for instance, continues to have a positive effect on the probability of being chosen from the queue and a negative effect on the probability of queuing for formal jobs. The major differences are in the "in the queue" equation, where the coefficient on tenure turns out to be negative and significant[62], and the coefficient on sex becomes positive and significant.

Overall, the estimation of different specifications for the structural bivariate probit seems to indicate that the wage differential is, indeed, a major contributor for the decision to queue for formal jobs. Similarly, schooling seems to be the main screening device used by formal employers when selecting workers from the queue. The most remarkable difference between the sequential decision specifications refers to the effect of have been an informal worker in the last job on both equations. The univariate probit specification indicates a negative impact of being a "former informal worker", i.e, workers who worked in the informal sector in their last job are less likely to be in the formal sector in the current one. However, whereas the structural Abowd-Farber bivariate probit shows that this negative effect is due to the fact that former informal sector workers are less likely to join the queue for formal jobs, the structural bivariate probit with sample selection indicates that this negative effect is due to employers "discriminating" against former informal workers.

[60] As we cannot estimate a "selectivity-corrected" wage equation for informal workers in the case of a bivariate probit with simultaneous decision, we do not estimate the structural bivariate probit for this model.

[61] Notice that the marginal effects of the wage differential in the "in the queue" equation and of the formal sector wage in the "chosen from the queue" equation is smaller in this case than in the Abowd-Faber bivariate probit.

[62] Notice that only in this specification we could find a negative sign for seniority as expected according to the theory. Workers with seniority should be less likely to queue since they would lose the benefits that accrue from it.

	BP with sample selection (A)				BP with sample selection (B)			
	IQ (In the queue=1)		CFQ (formal sector=1)		IQ		CFQ	
	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
(W_f - W_i)	19.38	0.467			20.52	0.493		
($W_f \| IQ = 1$)			0.08	0.039				
Constant	-8.95	0.259	-1.18	0.047	-9.28	0.268	-1.24	0.036
Some elementary	-0.09	0.110	0.19	0.028	-0.16	0.111	0.19	0.028
Complete elementary	0.03	0.136	0.52	0.028	-0.14	0.141	0.54	0.026
Complete primary	-1.64	0.215	1.03	0.037	-1.99	0.220	1.07	0.031
Complete secondary	-3.52	0.370	1.48	0.052	-4.36	0.457	1.56	0.033
Complete college	-3.37	0.350	1.68	0.090	-3.41	0.363	1.82	0.052
Exp/10	0.65	0.113	0.45	0.025	0.53	0.116	0.47	0.022
$Exp^2/100$	0.02	0.018	-0.05	0.004	0.04	0.018	-0.05	0.004
Tenure	-0.32	0.011			-0.33	0.011		
Sex (male=1)	1.25	0.080	0.18	0.020	1.25	0.081	0.21	0.014
Metropolitan	1.78	0.091	0.22	0.014	1.83	0.095	0.23	0.014
Race (white=1)	-1.25	0.101	0.09	0.016	-1.46	0.105	0.10	0.015
More than 5 years	0.06	0.136	0.47	0.024	0.16	0.139	0.47	0.024
Former formal worker	0.50	0.137	0.33	0.021	0.47	0.143	0.32	0.020
Former informal worker	0.22	0.092	-0.47	0.020	0.30	0.093	-0.47	0.020
Former public servant	0.37	0.502	0.16	0.102	0.20	0.530	0.22	0.101
Former self-employed	0.11	0.180	-0.22	0.035	0.16	0.182	-0.23	0.034
#Children 0-3 years	0.05	0.048			0.03	0.049		
#Children 4-6 years	-0.03	0.085			0.01	0.086		
#Children7-10 years	-0.17	0.062			-0.13	0.062		
#Children 11-13 year	0.07	0.052			0.06	0.052		
# Elder	0.01	0.108			-0.03	0.112		
# Formal	0.21	0.041			0.17	0.042		
#Self-employed	-0.02	0.070			-0.03	0.071		
#Informal	-0.09	0.033			-0.10	0.033		
#Unemployed	0.11	0.127			0.12	0.131		
Per capita income (others)	-0.33	0.127			-0.22	0.209		
Married	0.86	0.115	0.15	0.018	0.78	0.118	0.16	0.017
North	1.18	0.846	0.15	0.030	1.51	0.720	0.19	0.026
Southeast	0.14	0.127	0.42	0.021	0.13	0.133	0.44	0.018
South	1.09	0.191	0.54	0.026	1.17	0.198	0.56	0.024
Mid-West	0.79	0.109	0.12	0.028	0.89	0.111	0.16	0.024
Log L	-22924				-22917			
N	60138							

Table 3.12: Bivariate Probit with Sample Selection Structural Bivariate Probit Specification

3.6 Robustness Checks

3.6.1 Tenure as an Exclusion Restriction in the Bivariate Probit Models

Following the literature[63] we have excluded tenure (seniority) from the "chosen from the queue" equation of the bivariate probit models estimated in this paper. The argument is that as seniority is basically "specific human capital", the new employer should not look at it when deciding to hire or not an employee. Much more important for the employer's decision would be the worker's experience in the labour market[64]. Besides, as we included in both equations the employment history of the worker, we

[63]See Farber (1983) and Abowd and Farber (1982).

[64]Notice that the variable experience is not "true experience", but the traditional proxy: $age - yearsofeducation - 6$.

	Without tenure in the CFQ				With tenure in the CFQ			
	IQ		CFQ		IQ		CFQ	
dep. var: formal=1	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Constant	0.32	0.080	-0.44	0.064	-0.60	0.061	0.71	0.091
Some elementary	-0.06	0.054	0.19	0.044	0.15	0.041	-0.04	0.062
Complete elementary	-0.01	0.052	0.46	0.042	0.41	0.039	-0.03	0.059
Complete primary	0.02	0.060	0.93	0.050	0.88	0.047	-0.08	0.067
Complete secondary	0.08	0.063	1.30	0.052	1.23	0.050	-0.04	0.069
Complete college	0.06	0.077	1.26	0.068	1.28	0.066	-0.14	0.084
Exp/10	0.09	0.037	0.49	0.035	0.62	0.034	-0.11	0.042
$Exp^2/100$	-0.04	0.007	-0.06	0.006	-0.08	0.006	-0.01	0.007
Tenure	0.40	0.012			-0.003	0.002	0.51	0.016
Sex (male=1)	0.00	0.026	-0.01	0.027	0.01	0.024	-0.01	0.030
Metropolitan	0.16	0.025	0.25	0.024	0.22	0.022	0.16	0.028
Race (white=1)	-0.02	0.026	0.03	0.025	0.03	0.023	-0.03	0.029
More than 5 years	-0.45	0.057	0.28	0.030	0.28	0.028	-0.53	0.065
Former formal worker	0.27	0.044	1.02	0.048	0.86	0.039	0.24	0.052
Former informal worker	-0.41	0.050	0.33	0.057	0.17	0.043	-0.45	0.058
Former public servant	-0.12	0.123	1.10	0.415	9.62	1.4E+12	-0.39	0.124
Former self-employed	-0.31	0.060	0.87	0.126	0.53	0.082	-0.31	0.070
#Children 0-3 years	-0.02	0.015			-0.04	0.013		
#Children 4-6 years	-0.01	0.026			0.03	0.023		
#Children7-10 years	-0.04	0.019			-0.02	0.016		
#Children 11-13 year	-0.01	0.018			-0.02	0.015		
#Elder	-0.09	0.039			-0.06	0.033		
# Formal	0.25	0.014			0.25	0.011		
#Self-employed	-0.08	0.019			-0.09	0.016		
#Informal	-0.23	0.012			-0.23	0.011		
#Unemployed	0.03	0.026			0.07	0.025		
Per capita income (others)	-0.21	0.030			-0.17	0.029		
Married	0.04	0.028	0.14	0.027	0.17	0.025	-0.01	0.030
North	0.07	0.047	0.07	0.041	0.04	0.038	0.09	0.055
Southeast	0.19	0.033	0.34	0.031	0.29	0.028	0.19	0.037
South	0.32	0.041	0.37	0.038	0.32	0.035	0.34	0.046
Mid-West	0.03	0.041	0.08	0.040	0.08	0.036	0.02	0.047
Constructing	0.24	0.068	-1.11	0.044	-0.96	0.040	0.27	0.075
Retail	-0.08	0.036	-0.37	0.038	-0.30	0.034	-0.12	0.041
Lodging, food and services	-0.48	0.040	-1.10	0.036	-1.02	0.032	-0.48	0.044
Productive Sector	-0.13	0.032	-0.07	0.037	-0.03	0.033	-0.20	0.037
Social Services	-0.27	0.040	-0.07	0.045	-0.02	0.041	-0.37	0.045
Log L	-25059				-24856			
N	60138				60138			

Table 3.13: Sensitivity of the Abowd-Faber Bivariate Probit to the Inclusion of Tenure in the CFQ Equation

are already assessing the effect of the kind of experience that the individual had in the labour market[65], an information that could be used as a screening device by the employers.

Nevertheless, unlike the other exclusion restrictions to identify the "chosen from the queue" equation that are related to the situation of the household where the individual resides, tenure is related to the labour market and as such can also determine the employer's decision. For that reason we re-estimate the bivariate probit models including tenure as an additional regressor in the "chosen from the queue" equation.

[65]Actually the dummy for "more than 5 years in the current job" - one of the variables related to the employment history - is already a proxy for seniority.

	Without tenure in the CFQ				With tenure in the CFQ			
	IQ		CFQ		IQ		CFQ	
formal sector=1	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Some elementary	-0.01	*0.031*	0.17	*0.026*	-0.01	*0.031*	0.15	*0.026*
Complete elementary	-0.05	*0.030*	0.45	*0.025*	-0.04	*0.030*	0.40	*0.025*
Complete primary	0.00	*0.036*	0.86	*0.030*	0.01	*0.036*	0.77	*0.030*
Complete secondary	0.06	*0.038*	1.26	*0.033*	0.08	*0.038*	1.13	*0.033*
Complete college	0.05	*0.050*	1.51	*0.052*	0.08	*0.050*	1.32	*0.053*
Exp/10	0.08	*0.024*	0.46	*0.021*	0.10	*0.024*	0.35	*0.022*
$Exp^2/100$	-0.04	*0.004*	-0.05	*0.004*	-0.04	*0.004*	-0.04	*0.004*
Tenure	0.03	*0.002*			0.02	*0.002*	0.06	*0.002*
Sex (male=1)	-0.08	*0.018*	0.12	*0.015*	-0.07	*0.018*	0.09	*0.016*
Metropolitan	0.08	*0.017*	0.23	*0.014*	0.07	*0.017*	0.23	*0.014*
Race (white=1)	-0.08	*0.018*	0.09	*0.014*	-0.08	*0.018*	0.09	*0.015*
More than 5 years	0.20	*0.025*	0.41	*0.024*	0.19	*0.025*	0.26	*0.024*
Former formal worker	0.51	*0.025*	0.30	*0.020*	0.47	*0.025*	0.49	*0.021*
Former informal worker	0.21	*0.024*	-0.36	*0.019*	0.16	*0.024*	-0.19	*0.020*
Former public servant	0.13	*0.111*	0.19	*0.097*	0.08	*0.110*	0.36	*0.100*
Former self-employed	0.13	*0.039*	-0.14	*0.032*	0.07	*0.039*	0.08	*0.033*
#Children 0-3 years	-0.02	*0.011*			-0.02	*0.011*		
#Children 4-6 years	0.03	*0.019*			0.03	*0.019*		
#Children7-10 years	-0.01	*0.014*			-0.01	*0.014*		
#Children 11-13 year	0.02	*0.013*			0.02	*0.013*		
# Elder	0.00	*0.030*			0.00	*0.030*		
# Formal	0.16	*0.009*			0.16	*0.009*		
#Self-employed	-0.03	*0.014*			-0.03	*0.014*		
#Informal	-0.09	*0.009*			-0.09	*0.009*		
#Unemployed	0.08	*0.022*			0.08	*0.022*		
Per capita income (others)	-0.27	*0.020*			-0.27	*0.019*		
Married	0.03	*0.019*	0.12	*0.016*	0.03	*0.019*	0.12	*0.016*
North	-0.22	*0.030*	0.20	*0.025*	-0.23	*0.029*	0.22	*0.025*
Southeast	-0.09	*0.023*	0.41	*0.018*	-0.09	*0.023*	0.41	*0.018*
South	-0.07	*0.029*	0.52	*0.024*	-0.07	*0.029*	0.52	*0.024*
Mid-West	-0.21	*0.026*	0.22	*0.022*	-0.21	*0.026*	0.22	*0.022*
Constructing	-0.13	*0.032*	-0.67	*0.024*	-0.12	*0.032*	-0.64	*0.024*
Retail	-0.21	*0.027*	-0.18	*0.022*	-0.21	*0.027*	-0.17	*0.022*
Lodging, food and services	-0.44	*0.024*	-0.75	*0.020*	-0.44	*0.024*	-0.73	*0.020*
Productive Sector	0.00	*0.026*	-0.10	*0.022*	0.00	*0.026*	-0.11	*0.023*
Social Services	-0.09	*0.032*	-0.14	*0.028*	-0.09	*0.032*	-0.14	*0.028*
$\rho_{1,2}$	-0.96	*0.011*			-0.9618	*0.0117*		
Log L	-35326				-34950			
N	60138				60138			

Table 3.14: Sensitivity of the Poirier Bivariate Probit to the inclusion of Tenure in the CFQ

	Without tenure in the CFQ				With tenure in the CFQ			
	IQ (In the queue=1)		CFQ (formal sector=1)		IQ		CFQ	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Constant	0.71	*0.082*	-0.42	*0.062*	-0.53	*0.060*	0.99	*0.091*
Some elementary	-0.10	*0.055*	0.19	*0.043*	0.16	*0.040*	-0.09	*0.061*
Complete elementary	-0.10	*0.053*	0.44	*0.040*	0.41	*0.038*	-0.13	*0.059*
Complete primary	-0.13	*0.062*	0.89	*0.048*	0.85	*0.046*	-0.22	*0.068*
Complete secondary	-0.12	*0.065*	1.28	*0.051*	1.20	*0.048*	-0.21	*0.070*
Complete college	-0.10	*0.079*	1.26	*0.066*	1.26	*0.065*	-0.29	*0.085*
Exp/10	0.03	*0.038*	0.47	*0.034*	0.57	*0.033*	-0.15	*0.042*
$Exp^2/100$	-0.03	*0.007*	-0.06	*0.006*	-0.07	*0.006*	0.00	*0.007*
Tenure	0.39	*0.011*			0.00	*0.002*	0.51	*0.016*
Sex (male=1)	0.00	*0.028*	0.03	*0.025*	0.02	*0.024*	0.01	*0.030*
Metropolitan	0.11	*0.026*	0.24	*0.022*	0.21	*0.021*	0.12	*0.029*
Race (white=1)	-0.02	*0.027*	0.03	*0.024*	0.03	*0.022*	-0.03	*0.029*
More than 5 years	-0.46	*0.055*	0.27	*0.029*	0.27	*0.028*	-0.50	*0.062*
Former formal worker	0.16	*0.044*	0.89	*0.041*	0.77	*0.036*	0.18	*0.051*
Former informal worker	-0.45	*0.053*	0.36	*0.052*	0.22	*0.044*	-0.43	*0.058*
Former public servant	-0.24	*0.129*	0.99	*0.314*	0.76	*0.218*	-0.25	*0.140*
Former self-employed	-0.42	*0.063*	0.78	*0.092*	0.52	*0.073*	-0.35	*0.072*
#Children 0-3 years	-0.02	*0.014*			-0.03	*0.011*		
#Children 4-6 years	-0.02	*0.023*			0.02	*0.020*		
#Children7-10 years	-0.04	*0.017*			-0.02	*0.014*		
#Children 11-13 year	-0.01	*0.016*			-0.02	*0.013*		
# Elder	-0.09	*0.036*			-0.07	*0.029*		
# Formal	0.24	*0.013*			0.22	*0.010*		
#Self-employed	-0.08	*0.017*			-0.08	*0.014*		
#Informal	-0.20	*0.011*			-0.19	*0.009*		
#Unemployed	0.03	*0.024*			0.06	*0.021*		
Per capita income (others)	-0.19	*0.028*			-0.15	*0.026*		
Married	0.02	*0.029*	0.13	*0.025*	0.16	*0.024*	-0.03	*0.031*
North	0.05	*0.048*	0.08	*0.040*	0.03	*0.037*	0.10	*0.055*
Southeast	0.13	*0.034*	0.31	*0.029*	0.26	*0.028*	0.15	*0.038*
South	0.26	*0.043*	0.34	*0.036*	0.29	*0.034*	0.30	*0.047*
Mid-West	0.02	*0.043*	0.10	*0.038*	0.07	*0.035*	0.03	*0.048*
Constructing	0.49	*0.070*	-1.08	*0.043*	-0.95	*0.040*	0.52	*0.078*
Retail	-0.03	*0.038*	-0.33	*0.037*	-0.30	*0.033*	-0.05	*0.043*
Lodging, food and services	-0.18	*0.046*	-1.03	*0.035*	-0.96	*0.032*	-0.20	*0.048*
Productive Sector	-0.12	*0.034*	-0.05	*0.036*	-0.02	*0.033*	-0.19	*0.038*
Social Services	-0.26	*0.042*	-0.05	*0.044*	-0.02	*0.040*	-0.34	*0.046*
$\rho_{1,2}$	-0.80	*0.029*			-0.795	*0.0311*		
Log L	-24979				-24773			
N	60138				60138			

Table 3.15: Sensitivity of the Bivariate Probit with Sample Selection to the Inclusion of Tenure in the CFQ equation

The results of the Abowd-Farber bivariate probit with partial observability and of the Poirier biraviate probit with partial observability are very different when we include tenure (See Table 3.13 and Table 3.14). The sign of the variables related to the past experience change in both equations and so do the coefficients of the variables related to education[66]. According to the new results, education seems to increase the probability of being in the queue and does not have any effect on the probability of being chosen from the queue. The change in the variables related to the individual labour market history does not come as a surprise since tenure must be correlated with them, but the change in the signal of the education dummies are quite surprising. One possible explanation for that is the fact that both Abowd-Faber bivariate probit and Poirier bivariate probit with partial observability rely heavily on the exclusion restrictions and on the non-linearities of the model in order to identify separately the two equations. As for the impact of tenure on the two equations, it has no effect on the probability of being in the queue and increases the probability of being chosen from the queue.

Table 3.15 shows the results for the bivariate probit with sample selection. Unlike the bivariate probit models with partial observability, the bivariate probit with sample selection shows much more stability on their estimates regardless of the inclusion of tenure as an explanatory variable in the "chosen from the queue" equation. The only variables whose coefficients magnitude change, but not the sign, are those related to the labour market history of the individual as should be expected given their correlation with tenure. As for its effect, tenure seems to affect positively both the probability of being in the queue and the probability of being "chosen from the queue".

These results mean that the bivariate probit with sample selection seems to be the most suitable model to treat the "job queue" issue, not only because it offers a richer specification than the other two bivariate probit with partial observability, but also because it is much less sensitive to changes in the exclusion restrictions[67].

[66]But only the coefficients of the variables related to labour market history are statistically significant.

[67]As for the impact in the wage equations corrected for selectivity, we do not observe any major changes in the estimated parameters due to the inclusion of tenure in the "chosen from the queue". The tables with these results are available upon request.

3.6.2 The Stability of the Parameters of the Wage Equations

Tables 3.16 to Tables 3.19 and 3.20 compare the different wage equation estimates that we get if we exclude one set of exclusion restriction per time from our previous estimations[68]. We report the full model, and then we report the results with the variables omitted. These variables are: 1) the variables related to the number of children and older people in the household; 2) the past employment variables; 3) other member of household income (per capita); 4) variable related to the occupational distribution of the other members of the household.

[68] An alternative way of testing the robustness of those results was to use IV estimator and Hausman test for model misspecification. However, it is not clear how this sort of procedure could be adapted to a bivariate sample selection as the one developed here.

	OLS - 2STEP											
	Full Model				Excluding #children and #elder				Excluding past employment			
	Formal		Informal		Formal		Informal		Formal		Informal	
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-0.96	0.03	-1.80	0.03	-0.96	0.03	-1.80	0.03	-1.10	0.03	-1.78	0.03
Some elementary	0.14	0.02	0.15	0.02	0.14	0.02	0.15	0.02	0.15	0.02	0.16	0.02
Complete elementary	0.32	0.01	0.36	0.02	0.33	0.01	0.36	0.02	0.35	0.02	0.38	0.02
Complete primary	0.64	0.02	0.64	0.03	0.64	0.02	0.65	0.03	0.69	0.02	0.69	0.03
Complete secondary	1.12	0.02	1.07	0.03	1.12	0.02	1.07	0.03	1.18	0.02	1.13	0.04
Complete college	1.93	0.02	1.89	0.05	1.93	0.02	1.89	0.05	1.99	0.02	1.96	0.05
Exp/10	0.30	0.01	0.36	0.02	0.30	0.01	0.37	0.02	0.32	0.01	0.39	0.02
$Exp^2/100$	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.06	0.00	-0.05	0.00	-0.06	0.00
Tenure	0.03	0.00	0.01	0.00	0.03	0.00	0.01	0.00	0.03	0.00	0.01	0.00
Sex (male=1)	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01
Metropolitan	0.09	0.01	0.19	0.01	0.09	0.01	0.19	0.01	0.10	0.01	0.20	0.01
Race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
Married	0.17	0.01	0.21	0.01	0.17	0.01	0.21	0.01	0.18	0.01	0.22	0.01
North	0.42	0.01	0.48	0.02	0.42	0.01	0.48	0.02	0.43	0.01	0.48	0.02
Southeast	0.28	0.01	0.31	0.02	0.28	0.01	0.31	0.02	0.30	0.01	0.33	0.02
South	0.27	0.01	0.35	0.02	0.27	0.01	0.36	0.02	0.30	0.01	0.38	0.02
Mid-West	0.41	0.01	0.46	0.02	0.41	0.01	0.46	0.02	0.41	0.01	0.46	0.02
Constructing	-0.01	0.01	0.14	0.02	-0.01	0.01	0.14	0.02	-0.05	0.01	0.10	0.02
Retail	-0.21	0.01	-0.02	0.02	-0.21	0.01	-0.03	0.02	-0.23	0.01	-0.04	0.02
Lodging, food and services	-0.22	0.01	-0.13	0.02	-0.22	0.01	-0.14	0.02	-0.28	0.02	-0.18	0.02
Productive Sector	0.02	0.01	0.10	0.02	0.02	0.01	0.10	0.02	0.02	0.01	0.10	0.02
Social Services	-0.16	0.01	-0.06	0.02	-0.16	0.01	-0.06	0.02	-0.17	0.01	-0.07	0.02
Inverse Mills ratio (lambda)	-0.20	0.02	-0.21	0.02	-0.20	0.02	-0.20	0.02	-0.05	0.02	-0.12	0.03
	Excluding other's income				excluding other's ocupation							
	Formal		Informal		Formal		Informal					
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.				
Constant	-0.90	0.03	-1.81	0.03	-0.98	0.03	-1.82	0.03				
Some elementary	0.14	0.02	0.14	0.02	0.14	0.02	0.15	0.02				
Complete elementary	0.31	0.01	0.35	0.02	0.33	0.02	0.36	0.02				
Complete primary	0.62	0.02	0.62	0.03	0.65	0.02	0.64	0.03				
Complete secondary	1.10	0.02	1.02	0.03	1.13	0.02	1.05	0.03				
Complete college	1.91	0.02	1.84	0.05	1.94	0.02	1.88	0.05				
Exp/10	0.29	0.01	0.35	0.02	0.30	0.01	0.36	0.02				
$Exp^2/100$	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.05	0.00				
Tenure	0.02	0.00	0.01	0.00	0.03	0.00	0.01	0.00				
Sex (male=1)	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01				
Metropolitan	0.08	0.01	0.18	0.01	0.09	0.01	0.18	0.01				
Race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01				
Married	0.16	0.01	0.20	0.01	0.17	0.01	0.21	0.01				
North	0.42	0.01	0.48	0.02	0.42	0.01	0.48	0.02				
Southeast	0.27	0.01	0.29	0.02	0.28	0.01	0.30	0.02				
South	0.26	0.01	0.34	0.02	0.28	0.01	0.35	0.02				
Mid-West	0.41	0.01	0.45	0.02	0.41	0.01	0.46	0.02				
Constructing	0.01	0.01	0.16	0.02	-0.02	0.01	0.15	0.02				
Retail	-0.21	0.01	-0.02	0.02	-0.22	0.01	-0.02	0.02				
Lodging, food and services	-0.19	0.01	-0.10	0.02	-0.24	0.01	-0.12	0.02				
Productive Sector	0.03	0.01	0.10	0.02	0.02	0.01	0.11	0.02				
Social Services	-0.16	0.01	-0.06	0.02	-0.17	0.01	-0.05	0.02				
Inverse Mills ratio (lambda)	-0.26	0.02	-0.27	0.02	-0.16	0.02	-0.23	0.03				

Table 3.16: OLS Wage Equations Based on the Univariate Probit with Different Exclusion Restrictions

	Maximum Likelihood Estimation											
	Full Model				Excluding #children and #elder				Excluding past employment			
	Formal		Informal		Formal		Informal		Formal		Informal	
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-0.83	*0.023*	-1.80	*0.027*	-0.83	0.02	-1.80	0.03	-0.82	0.02	-1.80	0.03
Some elementary	0.13	*0.017*	0.15	*0.019*	0.13	0.02	0.15	0.02	0.13	0.02	0.15	0.02
Complete elementary	0.30	*0.016*	0.36	*0.019*	0.30	0.02	0.36	0.02	0.30	0.02	0.37	0.02
Complete primary	0.60	*0.018*	0.63	*0.025*	0.60	0.02	0.64	0.03	0.59	0.02	0.65	0.03
Complete secondary	1.06	*0.018*	1.05	*0.030*	1.06	0.02	1.05	0.03	1.06	0.02	1.08	0.03
Complete college	1.88	*0.020*	1.86	*0.040*	1.88	0.02	1.87	0.04	1.88	0.02	1.89	0.04
Exp/10	0.27	*0.010*	0.36	*0.017*	0.27	0.01	0.36	0.02	0.28	0.01	0.37	0.02
$Exp^2/100$	-0.05	*0.002*	-0.05	*0.003*	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.06	0.00
Tenure	0.02	*0.001*	0.01	*0.001*	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00
Sex (male=1)	0.30	*0.007*	0.38	*0.012*	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01
Metropolitan	0.08	*0.006*	0.18	*0.012*	0.08	0.01	0.18	0.01	0.07	0.01	0.19	0.01
Race (white=1)	0.16	*0.007*	0.11	*0.012*	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
Married	0.16	*0.007*	0.21	*0.013*	0.16	0.01	0.21	0.01	0.16	0.01	0.21	0.01
North	0.42	*0.011*	0.47	*0.018*	0.42	0.01	0.47	0.02	0.42	0.01	0.48	0.02
Southeast	0.26	*0.009*	0.30	*0.015*	0.26	0.01	0.31	0.02	0.26	0.01	0.31	0.02
South	0.25	*0.011*	0.35	*0.022*	0.24	0.01	0.35	0.02	0.24	0.01	0.36	0.02
Mid-West	0.41	*0.011*	0.46	*0.017*	0.41	0.01	0.46	0.02	0.41	0.01	0.46	0.02
Constructing	0.03	*0.014*	0.14	*0.022*	0.03	0.01	0.14	0.02	0.03	0.01	0.13	0.02
Retail	-0.20	*0.010*	-0.03	*0.019*	-0.20	0.01	-0.03	0.02	-0.20	0.01	-0.03	0.02
Lodging, food and services	-0.16	*0.012*	-0.13	*0.020*	-0.16	0.01	-0.13	0.02	-0.15	0.01	-0.15	0.02
Productive Sector	0.03	*0.008*	0.10	*0.019*	0.03	0.01	0.10	0.02	0.03	0.01	0.10	0.02
Social Services	-0.15	*0.011*	-0.06	*0.024*	-0.15	0.01	-0.06	0.02	-0.16	0.01	-0.06	0.02
$\rho_{1,u}$	-0.53	*0.015*	-0.33	*0.028*	-0.53		-0.32		-0.52		-0.27	
	Excluding other's income				Excluding other's ocupation							
	Formal		Informal		Formal		Informal					
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.				
Constant	-0.85	0.02	-1.80	0.03	-0.81	0.02	-1.82	0.03				
some elementary	0.13	0.02	0.15	0.02	0.13	0.02	0.15	0.02				
complete elementary	0.30	0.02	0.36	0.02	0.30	0.02	0.35	0.02				
complete primary	0.61	0.02	0.65	0.03	0.59	0.02	0.62	0.03				
complete secondary	1.07	0.02	1.07	0.03	1.06	0.02	1.04	0.03				
complete college	1.89	0.02	1.89	0.04	1.87	0.02	1.85	0.04				
exp/10	0.28	0.01	0.37	0.02	0.27	0.01	0.36	0.02				
$Exp^2/100$	-0.05	0.00	-0.06	0.00	-0.05	0.00	-0.05	0.00				
tenure	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00				
sex (male=1)	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01				
metropolitan	0.08	0.01	0.19	0.01	0.07	0.01	0.18	0.01				
race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01				
married	0.16	0.01	0.21	0.01	0.16	0.01	0.21	0.01				
North	0.42	0.01	0.48	0.02	0.42	0.01	0.47	0.02				
Southeast	0.26	0.01	0.31	0.02	0.26	0.01	0.30	0.02				
South	0.25	0.01	0.36	0.02	0.24	0.01	0.35	0.02				
Mid-West	0.41	0.01	0.46	0.02	0.40	0.01	0.46	0.02				
constructing	0.02	0.01	0.13	0.02	0.03	0.01	0.16	0.02				
retail	-0.20	0.01	-0.03	0.02	-0.20	0.01	-0.02	0.02				
lodging, food and services	-0.17	0.01	-0.14	0.02	-0.15	0.01	-0.11	0.02				
Productive Sector	0.03	0.01	0.10	0.02	0.03	0.01	0.11	0.02				
Social Services	-0.16	0.01	-0.07	0.02	-0.15	0.01	-0.06	0.02				
$\rho_{1,u}$	-0.49		-0.29		-0.35	0.03	-0.53	0.02				

Table 3.17: MLE Wage Equations Based on the Univariate Probit with Different Exclusion Restrictions

	ABOWD-FABER (A)											
	Full Model				Excluding #children and #elder				Excluding past employment			
	Formal		Informal		Formal		Informal		Formal		Informal	
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-1.02	0.03	-1.69	0.04	-1.02	0.03	-1.69	0.04	-1.55	0.04	-1.68	0.09
Some elementary	0.15	0.02	0.14	0.02	0.15	0.02	0.14	0.02	0.20	0.02	0.17	0.02
Complete elementary	0.34	0.02	0.33	0.02	0.34	0.01	0.33	0.02	0.46	0.02	0.38	0.04
Complete primary	0.67	0.02	0.57	0.03	0.67	0.02	0.58	0.03	0.87	0.02	0.69	0.08
Complete secondary	1.16	0.02	0.97	0.04	1.16	0.02	0.97	0.04	1.42	0.03	1.15	0.12
Complete college	1.97	0.02	1.81	0.05	1.97	0.02	1.81	0.05	2.21	0.03	1.98	0.12
Exp/10	0.31	0.01	0.32	0.02	0.31	0.01	0.32	0.02	0.42	0.01	0.39	0.05
$Exp^2/100$	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.07	0.00	-0.06	0.01
Tenure	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00	0.02	0.00	0.00	0.00
Sex (male=1)	0.30	0.01	0.36	0.01	0.30	0.01	0.37	0.01	0.30	0.01	0.37	0.01
Metropolitan	0.10	0.01	0.16	0.01	0.10	0.01	0.17	0.01	0.15	0.01	0.20	0.03
Race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
Married	0.17	0.01	0.20	0.01	0.17	0.01	0.20	0.01	0.21	0.01	0.22	0.02
North	0.42	0.01	0.47	0.02	0.42	0.01	0.47	0.02	0.42	0.01	0.47	0.02
Southeast	0.29	0.01	0.28	0.02	0.29	0.01	0.28	0.02	0.37	0.01	0.33	0.04
South	0.28	0.01	0.32	0.02	0.28	0.01	0.33	0.02	0.38	0.01	0.38	0.04
Mid-West	0.41	0.01	0.45	0.02	0.41	0.01	0.45	0.02	0.42	0.01	0.46	0.02
constructing	-0.04	0.02	0.19	0.03	-0.04	0.02	0.19	0.03	-0.28	0.02	0.06	0.09
Retail	-0.22	0.01	0.00	0.02	-0.22	0.01	0.00	0.02	-0.29	0.01	-0.05	0.04
Lodging, food and services	-0.25	0.01	-0.06	0.02	-0.25	0.01	-0.06	0.02	-0.55	0.03	-0.20	0.10
Productive Sector	0.02	0.01	0.11	0.02	0.02	0.01	0.11	0.02	0.00	0.01	0.11	0.02
Social Services	-0.17	0.01	-0.05	0.02	-0.17	0.01	-0.05	0.02	-0.20	0.01	-0.06	0.03
σ_{1f}	-0.19	0.02			-0.18	0.02			-0.13	0.02		
σ_{2f}	-0.07	0.03			-0.07	0.03			0.63	0.06		
σ_{1i}			-0.18	0.02			-0.17	0.02			-0.14	0.02
δ_1			-0.18	0.02			-0.18	0.02			-0.11	0.02
δ_2			-0.18	0.02			-0.18	0.02			-0.04	0.13
R^2	0.56		0.47		0.56		0.47		0.56		0.46	

	Excluding other's income				excluding other's ocupation			
	Formal		Informal		Formal		Informal	
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-1.01	0.03	-1.69	0.04	-1.03	0.03	-1.73	0.04
Some elementary	0.15	0.02	0.14	0.02	0.15	0.02	0.14	0.02
Complete elementary	0.34	0.01	0.33	0.02	0.34	0.02	0.33	0.02
Complete primary	0.67	0.02	0.57	0.03	0.67	0.02	0.58	0.03
Complete secondary	1.16	0.02	0.97	0.04	1.16	0.02	0.98	0.04
Complete college	1.97	0.02	1.81	0.05	1.97	0.02	1.81	0.05
Exp/10	0.31	0.01	0.32	0.02	0.31	0.01	0.33	0.02
$Exp^2/100$	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.05	0.00
Tenure	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00
Sex (male=1)	0.30	0.01	0.36	0.01	0.30	0.01	0.37	0.01
Metropolitan	0.10	0.01	0.16	0.01	0.10	0.01	0.17	0.01
Race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
Married	0.17	0.01	0.20	0.01	0.17	0.01	0.20	0.01
North	0.42	0.01	0.47	0.02	0.42	0.01	0.47	0.02
Southeast	0.29	0.01	0.28	0.02	0.29	0.01	0.29	0.02
South	0.28	0.01	0.32	0.02	0.28	0.01	0.33	0.02
Mid-West	0.41	0.01	0.45	0.02	0.41	0.01	0.45	0.02
Constructing	-0.04	0.02	0.18	0.03	-0.04	0.01	0.19	0.03
Retail	-0.22	0.01	0.00	0.02	-0.22	0.01	0.00	0.02
Lodging, food and services	-0.25	0.01	-0.06	0.02	-0.26	0.01	-0.07	0.02
Productive Sector	0.02	0.01	0.11	0.02	0.02	0.01	0.11	0.02
Social Services	-0.17	0.01	-0.05	0.02	-0.17	0.01	-0.05	0.02
σ_{1f}	-0.22	0.02			-0.17	0.02		
σ_{2f}	-0.07	0.03			-0.08	0.02		
σ_{1i}			-0.21	0.02			-0.15	0.02
δ_1			-0.19	0.02			-0.15	0.03
δ_2			-0.18	0.02			-0.19	0.02
R^2	0.56		0.47		0.56		0.46	

Table 3.18: Wage Equations Based on the Abowd-Farber Bivariate Probit with Different Exclusion Restrictions

	Full Model						Excluding #children and #elder						Excluding past employment					
	Formal		Informal (IQ)		Informal (NIQ)		Formal		Informal (IQ)		Informal (NIQ)		Formal		Informal (IQ)		Informal (NIQ)	
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-0.99	0.029	-1.62	0.031	-1.62	0.115	-0.99	0.028	-1.62	0.031	-1.62	0.116	-1.23	0.037	-1.55	0.035	-1.01	0.141
Some elementary	0.14	0.016	0.11	0.021	0.21	0.042	0.14	0.015	0.11	0.020	0.21	0.042	0.17	0.016	0.19	0.028	0.22	0.042
Complete elementary	0.32	0.016	0.26	0.023	0.50	0.040	0.32	0.015	0.26	0.023	0.50	0.040	0.38	0.016	0.47	0.053	0.51	0.040
Complete primary	0.64	0.019	0.47	0.033	0.92	0.049	0.64	0.018	0.47	0.033	0.92	0.050	0.74	0.021	0.85	0.096	0.94	0.050
Complete secondary	1.11	0.020	0.81	0.044	1.37	0.052	1.11	0.020	0.81	0.044	1.37	0.052	1.24	0.024	1.37	0.141	1.41	0.052
Complete college	1.91	0.023	1.64	0.076	2.06	0.069	1.91	0.022	1.64	0.077	2.06	0.069	2.04	0.026	2.33	0.182	2.07	0.069
Exp/10	0.31	0.010	0.28	0.022	0.47	0.032	0.31	0.010	0.28	0.022	0.47	0.032	0.35	0.011	0.50	0.057	0.52	0.033
$Exp^2/100$	-0.06	0.002	-0.04	0.004	-0.07	0.006	-0.06	0.002	-0.04	0.004	-0.07	0.006	-0.06	0.002	-0.07	0.007	-0.09	0.006
Tenure	0.03	0.001	0.01	0.002	0.01	0.002	0.03	0.001	0.01	0.002	0.01	0.002	0.03	0.001	0.01	0.002	0.02	0.002
Sex (male=1)	0.29	0.007	0.35	0.014	0.35	0.025	0.29	0.007	0.35	0.014	0.35	0.025	0.30	0.007	0.41	0.020	0.35	0.025
Metropolitan	0.10	0.007	0.19	0.014	0.18	0.025	0.10	0.006	0.19	0.014	0.18	0.025	0.12	0.007	0.29	0.027	0.21	0.025
Race (white=1)	0.15	0.007	0.07	0.013	0.16	0.025	0.15	0.007	0.07	0.013	0.16	0.025	0.16	0.007	0.10	0.015	0.13	0.025
Married	0.17	0.007	0.17	0.015	0.30	0.026	0.17	0.007	0.17	0.015	0.30	0.026	0.18	0.007	0.23	0.020	0.31	0.026
North	0.40	0.013	0.40	0.021	0.53	0.042	0.40	0.013	0.40	0.021	0.53	0.042	0.42	0.013	0.47	0.026	0.46	0.043
Southeast	0.28	0.009	0.30	0.018	0.23	0.033	0.28	0.009	0.30	0.017	0.23	0.033	0.32	0.010	0.46	0.040	0.23	0.033
South	0.27	0.011	0.34	0.025	0.28	0.041	0.27	0.011	0.34	0.025	0.28	0.041	0.32	0.012	0.55	0.055	0.28	0.041
Mid-West	0.38	0.012	0.47	0.019	0.33	0.038	0.38	0.012	0.47	0.018	0.33	0.038	0.41	0.012	0.53	0.024	0.27	0.039
Constructing	-0.02	0.015	0.15	0.025	0.13	0.046	-0.02	0.014	0.15	0.025	0.13	0.046	-0.09	0.016	-0.10	0.066	0.09	0.046
Retail	-0.23	0.010	-0.08	0.021	0.04	0.040	-0.23	0.010	-0.08	0.021	0.03	0.040	-0.24	0.010	-0.15	0.027	-0.02	0.040
Lodging, food and services	-0.29	0.013	-0.15	0.024	-0.18	0.038	-0.29	0.012	-0.15	0.024	-0.18	0.038	-0.37	0.014	-0.44	0.072	-0.31	0.042
Productive Sector	0.02	0.009	0.06	0.022	0.20	0.039	0.02	0.008	0.06	0.022	0.20	0.039	0.01	0.008	0.03	0.024	0.20	0.039
Social Services	-0.18	0.011	-0.02	0.027	-0.15	0.047	-0.18	0.011	-0.02	0.028	-0.15	0.047	-0.18	0.011	-0.07	0.031	-0.18	0.047
σ_{1f}	0.31	0.037					0.30	0.037					0.13	0.044				
σ_{2f}	-0.24	0.030					-0.24	0.030					0.03	0.040				
σ_{1i}^{*}			-0.89	0.123					-0.89	0.114					-1.03	0.116		
σ_{1i}					-0.01	0.054					-0.01	0.054					0.32	0.070
σ_{2i}			-0.17	0.030					-0.17	0.030					0.34	0.124		
R^2	0.56		0.40		0.48		0.56		0.40		0.48		0.56		0.40		0.48	

Table 3.19: Wage Equations based on the Bivariate probit with Sample Selection with Different Exclusion Restrictions -1st part (cont.)

	Excluding other's income						Excluding other's ocupation					
	Formal		Informal (IQ)		Informal (NIQ)		Formal		Informal (IQ)		Informal (NIQ)	
dep. var: lwh	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
Constant	-1.10	0.03	-1.62	0.03	-2.21	0.12	-0.91	0.03	-1.63	0.03	-1.46	0.15
Some elementary	0.15	0.02	0.11	0.02	0.21	0.04	0.14	0.02	0.11	0.02	0.22	0.04
Complete elementary	0.35	0.02	0.26	0.02	0.50	0.04	0.30	0.02	0.27	0.02	0.50	0.04
Complete primary	0.69	0.02	0.47	0.03	0.89	0.05	0.61	0.02	0.48	0.03	0.92	0.05
Complete secondary	1.18	0.02	0.80	0.04	1.33	0.05	1.07	0.02	0.82	0.04	1.38	0.05
Complete college	1.99	0.02	1.64	0.08	2.05	0.07	1.86	0.02	1.66	0.08	2.06	0.07
Exp/10	0.32	0.01	0.28	0.02	0.43	0.03	0.30	0.01	0.29	0.02	0.48	0.03
$Exp^2/100$	-0.05	0.00	-0.04	0.00	-0.06	0.01	-0.06	0.00	-0.05	0.00	-0.08	0.01
Tenure	0.03	0.00	0.01	0.00	0.01	0.00	0.03	0.00	0.01	0.00	0.01	0.00
Sex (male=1)	0.30	0.01	0.35	0.01	0.36	0.02	0.28	0.01	0.35	0.01	0.35	0.03
Metropolitan	0.11	0.01	0.19	0.01	0.15	0.02	0.10	0.01	0.20	0.01	0.19	0.03
Race (white=1)	0.16	0.01	0.07	0.01	0.18	0.03	0.14	0.01	0.07	0.01	0.15	0.03
Married	0.18	0.01	0.17	0.02	0.28	0.03	0.16	0.01	0.17	0.02	0.30	0.03
North	0.42	0.01	0.40	0.02	0.58	0.04	0.37	0.01	0.40	0.02	0.51	0.04
Southeast	0.30	0.01	0.30	0.02	0.23	0.03	0.26	0.01	0.30	0.02	0.23	0.03
South	0.30	0.01	0.34	0.02	0.28	0.04	0.25	0.01	0.35	0.02	0.28	0.04
Mid-West	0.41	0.01	0.47	0.02	0.38	0.04	0.36	0.01	0.46	0.02	0.31	0.04
Constructing	-0.05	0.01	0.15	0.02	0.16	0.05	0.01	0.01	0.15	0.02	0.12	0.05
Retail	-0.23	0.01	-0.08	0.02	0.09	0.04	-0.24	0.01	-0.09	0.02	0.02	0.04
Lodging, food and services	-0.29	0.01	-0.14	0.02	-0.07	0.04	-0.30	0.01	-0.16	0.02	-0.22	0.04
Productive Sector	0.02	0.01	0.06	0.02	0.20	0.04	0.02	0.01	0.06	0.02	0.20	0.04
Social Services	-0.18	0.01	-0.02	0.03	-0.13	0.05	-0.18	0.01	-0.02	0.03	-0.16	0.05
σ_{1f}	0.00	0.04					0.58	0.05				
σ_{2f}	-0.04	0.03					-0.42	0.03				
σ^*_{1i}			-0.92	0.11					-1.09	0.20		
σ_{1i}					-0.33	0.06					0.08	0.07
σ_{2i}			-0.17	0.03					-0.16	0.03		
R_2	0.55		0.40		0.48		0.56		0.40		0.48	

Table 3.20: Wage Equations based on the Bivariate probit with Sample Selection with Different Exclusion Restrictions-2nd part

The results indicate a high degree of stability of the parameters of the wage equations. The estimates of the human capital variables do not vary much with changes in the exclusion restrictions of the sector allocation equations. The only exception refers to the variables related to past employment history. The estimated parameters are quite sensitive to their exclusion, and tend to be higher than the ones we get from the full model. As these variables are jointly significant in all estimates of the bivariate probit and display sensible effects in both the "in the queue" and the "chosen from the queue" equation, the models that exclude these variables may not have captured the precise nature of the sector allocation process. A possible explanation is that past employment history variables may reveal to the employer unobserved skill/productivity, making it very hard to justify its exclusion.

3.7 Conclusion

This chapter tackled the issue of segmentation between formal and informal sector in Brazil. Unlike most studies on the literature that emphasize the wage differential between the two types of workers or different wage structures, we concentrate our attention on the relationship between unobservables that determine the sector allocation process and unobservables that determine wages. This approach has been used to test the hypothesis of comparative advantage in the sector allocation process based on the worker's choice as discussed in Chapter 2. However, we depart from this approach, as we argue that this univariate approach would hide the true nature of the allocation process, which according to the evidence presented in this chapter is best depicted as a two-decision process: the decision of a worker to join the queue for a formal job and the decision of the employer to pick him/her up from the pool of workers in the queue.

We applied both Abowd and Farber (1982) and Mengistae (1998) tests for the existence of a job queue for formal jobs in Brazil and none of them were able to reject this hypothesis. We also present estimates of the job queue "length"[69] for selected groups and show that non-white (1.43), female (1.37), illiterate (1.74), "new entrants"(1.45) and former informal workers (2.11) are the groups with the lower probability of being chosen from the queue conditional on being in the queue[70]. This result is particularly strong for workers whose last job was in the informal sector, suggesting that a spell in the informal sector may jeopardize the worker's chance

[69]Job length is this case is defined as the inverse of the probability of getting a formal job once in the queue.

[70]The numbers between parenthesis correspond to the length of the queue.

of getting a formal job. Assuming that workers really would prefer and would be better off if they would get a formal job, then these groups should receive special attention of public policies to encourage formal employers to hire them.

The estimation of the "in the queue" and of the "chosen from the queue" equation separately also allowed us to uncover some relationships that were hidden by the univariate procedure in the sector allocation process. The different impact of education levels in the "in the queue" and "chosen from the queue" equations is a good example: whereas workers from low education groups seem to "join the queue" with a higher probability than more educated workers, the latter are much more likely to be chosen from it.

The results for the wage equation show the existence of different types of selectivity. But the way selectivity affects wages differs according to the assumed model. In the case of sequential models, unobservables that make workers more likely to be chosen from the queue make them earn less than expected wages in the formal sector. Thus it seems that workers less likely to be chosen to work in the formal sector are the ones who benefit more from this condition. On the other hand, workers less likely to be chosen from the queue are the ones who command higher wages in the informal sector.

As for the structural bivariate probit, the results indicate that wage differential plays the most important role in the decision to join the queue. However, the coefficient on the wage in the formal sector in the "chosen from the queue" equation displays an unexpected positive sign[71]. We attribute this puzzling result to the fact that the wage regressor seems to be capturing the effect of the human-capital related variables.

In terms of the methodology to estimating bivariate probit with partial observability, robustness checks indicated that the bivariate probit with sample selection is much less sensitive to minor changes in the specification than the other types of bivariate probit[72]. Therefore, the availability of the information about the worker's willingness to change to a formal job was a fundamental piece of information that allowed us to relax the severe partial observability of the Abowd-Farber and Poirier's bivariate probits, yielding much more information on the process of allocation of workers between sectors.

A shortcoming for this analysis is that we had to rely on data from 1990. There-

[71]Note that the fact that the higher the probability of being chosen from the queue the lower the wage indicates that employers somehow manage to minimize the cost in the hiring process.

[72]The robustness checks also revealed that the wage equation results are quite stable as long as the past employment history variables are included among the exclusion restrictions that identify the wage equation separately.

fore, it is hard to know whether these results hold for the late 1990's after the market-oriented reforms in general and the trade liberalization process, in particular. The fall in the wage gap between formal and informal sectors, for instance, may have led to a fall in the size of the queue for formal jobs.

Chapter 4

The Impact of Trade Liberalisation on the Informal Sector in Brazil

4.1 Introduction

A common feature of several middle-income developing countries in the late 1980's and early 1990's was the undertaking of several structural reforms, particularly trade liberalisation measures[1]. Many recent studies have tried to assess the impact of these reforms on the labour market of these countries. Basically, researchers have looked for evidence of any of the Hecksher-Ohlin/Stolper-Samuelson (HOS) framework's implications in developing country labour markets[2]. More specifically, they have looked for evidence that trade liberalisation has triggered the following sequence of events: 1) increases in the relative price of unskilled intensive products/firms/industries; 2) a positive effect of these price increases on the demand for

[1]Besides structural reforms, Brazil also witnessed a successful stabilization Plan in 1994 that may have had some effect on the composition of employment and on wage rates. This is so because the appreciation of the exchange rate due to the high interest rates was one of the major components of the stabilization strategy. For this reason, we control for industry-specific exchange rate in our empirical strategies. Moreover, Pontual et al. (2004) have shown that a depreciation of the exchange rate affects net employment growth by increasing job creation and hires, that tariffs have no effect on job or worker flows, whereas import penetration decrease job growth by increasing job destruction. These results suggest that the exchange rate have a very important role on job and worker flows, even after controlling for openness measures and sector specificities.

[2]Most of these studies have been motivated by the attempt to find the "reverse picture" of the effects of trade liberalisation on developed countries. A large literature has argued that globalisation or the increase in trade flows between developed and developing countries throughout the 1980's and 1990's can explain at least part of the increase in wage inequality in developed countries (See Wood (1997)). If this were the case, then one should also observe a lower inequality in developing countries as a result of the increase in trade flows.

unskilled workers; 3) a reduction in the wage premium of skilled workers, leading to lower wage inequality in these countries[3], and finally 4) an increase in the share of skilled workers in all sectors due to the increase in the relative price of unskilled workers. Besides this interest on the distributional consequences of trade liberalisation based on the HOS framework, researchers have also tried to measure the impact of trade liberalisation on employment and on the wage structure as a way to assess the importance of rent-sharing in the protected sectors[4].

Despite these numerous studies on the impact of trade liberalisation on developing country labour markets, several questions remain to be explored. In particular, as pointed out by Behrman (1999), the impact of trade liberalisation on the "informal" manufacturing sector and the existence of possible spillover effects on the rest of the economy have been overlooked[5]. This is important because if dual labour markets are an important feature of developing countries, then to overlook the implications of trade liberalisation for the wage differential between "formal" and "informal" workers and on their mobility pattern[6] may yield an incomplete description of its impact on the entire labour market.

The aim of this chapter is to start filling this gap on the empirical literature using the Brazilian trade liberalisation experience as a quasi-natural experiment. We assess whether trade liberalisation can be considered a serious candidate to explain both the fall in the wage differential between registered (formal) and non-registered (informal) workers, and the fall in the proportion of registered workers in the economy. The episode of trade liberalisation in Brazil is particularly interesting for this task because it implied a huge fall in tariff and non-tariff barriers and it occurred during a short period of time, basically within 3 years. Besides, the schedule of tariff reduction announced in 1990 was brought forward several times, making a strong case for the exogeneity of the trade reforms[7].

[3]See for example Robbins and Gindling (1999) for Costa Rica, Robbins and Gindling (2001) for Costa Rica and Chile, Hanson and Harrison (1999) for Mexico, Beyer et al. (1999) for Chile, Green et al. (2001) and Gonzaga et al. (2002a) for Brazil.

[4]See for example Currie and Harrison (1997) for Morocco, Revenga (1997) for Mexico, Marquez and Pages (1997) for a panel of Latin American and Caribbean countries, Menezes-Filho and Arbache (2002) and Arbache and Corseuil (2003) for Brazil.

[5]Goldberg and Pavcnik (2003) is the first paper to the best of our knowledge to focus on this issue. They find that in Colombia trade liberalisation seems to have led to an increase in the manufacturing informal sector, but in Brazil there was no impact. They attribute such differences to the fact that the Brazilian labour market is more flexible (in terms of regulation) than the Colombian.

[6]Maloney (1997) associates some changes in the mobility pattern among Mexican workers, particularly the increase in the proportion of "contract" workers with the trade liberalisation in that country.

[7]Some commentators of the Brazilian trade liberalisation point out that the reform was launched much more as an auxiliary tool in the combat against the hyper-inflationary process witnessed by

The structure of the chapter is as follows. First we will describe the evolution of both the proportion of registered workers and the wage gap between registered and non-registered in the manufacturing sector during the 1980's and 1990's. Second, we will discuss the literature on the impact of the trade liberalisation on the labour market of developing countries, and the channels through which trade liberalisation could affect both the wage differential between registered and non-registered workers and the proportion of registered workers. Third, we will describe the main features of the trade liberalisation reform in Brazil. Fourth, we will put forward different procedures to identify whether or not trade liberalisation had any impact on the fall in the wage differential between registered and non-registered workers and on the increase of the proportion of non-registered workers. These procedures are based on: 1) exploiting industry variation of trade-related measures such as effective tariffs, import penetration and export orientation ratios for a panel of 17 tradable manufacturing industries; 2) exploiting regional variation of industry dispersion within the country, so that we will be able to test the impact, if any, of trade liberalisation on the entire labour market and not only in the manufacturing sector (spillover effect); 3) adding the non-tradable sector and cohort variation to the analysis. This latter approach allows us to use additional variation in industry-cohort cells in order to assess the impact of trade liberalisation on the tradable sector when compared to the non-tradable sector.

Our results suggest that trade liberalisation had a statistically significant impact on the reduction of the wage differential between registered and non-registered workers in the manufacturing sector. However, we do not find evidence of spillover to the entire economy. As for the impact on the proportion of registered workers, the results are not very robust, and in our opinion, it is not possible to make a strong case for the link between trade liberalisation and this phenomenon.

4.2 Registered and Non-registered Workers in the Manufacturing Sector in Brazil

As seen in chapter 2, the proportion of non-registered workers increased –at least, for non-farming activities – and the wage differential between the two groups diminished during the 1990's. As non-registered workers are less likely to be found in the manufacturing sector, one could associate the fall in the proportion of registered

the Brazilian economy in the early 1990's. See Muendler (2001) and Kume et al. (2000).

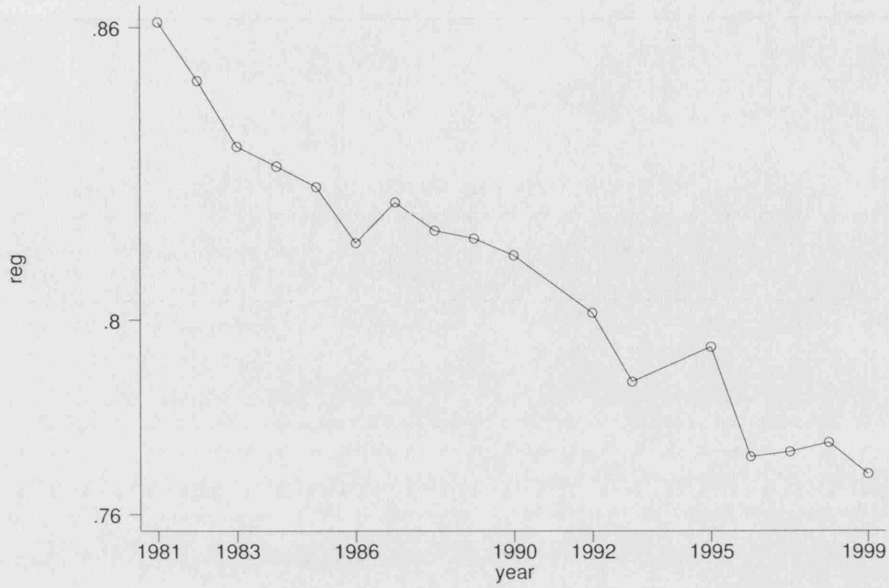

Figure 4.1: Proportion of Registered Workers in the Manufacturing Sector - 1981 - 1999

workers with the reduction in the number of workers in the manufacturing sector[8]. A lower proportion of "manufacturing jobs" would lead to a higher proportion of non-registered jobs in the whole economy. However, Figure 4.1 shows that the decrease in the proportion of registered workers was also observed in the manufacturing sector. From 86% in 1981, it was down to 77% in 1999.

As for the wage differential between registered and non-registered workers in the manufacturing sector[9], Figure 4.2 shows that it followed very closely the wage differential between registered and non-registered workers in the entire economy (See Figure 2.2 in Chapter 2)[10]. There were minor differences in the intensity of the movements, but their directions were basically the same. The wage differential between registered and non-registered in the manufacturing sector that had peaked at 86% in 1992 was down to 44% in 1999.

These figures reveal sharp movements in the time series of both the wage differential between registered and non-registered workers and the proportion of registered workers in the entire economy and in the manufacturing sector, in particular. Our

[8]The number of workers employed in the manufacturing sector peaked in 1989.

[9]The wage differential is calculated as $100 * [exp(b) - 1]$, where b is the coefficient of a dummy coefficient for registered workers in a standard semi-log wage equation, controlling for education (6 categories), experience, experience squared, gender, region, and metropolitan area.

[10]The only difference here is that the 1986 dip in the wage gap due to the *Cruzado Plan* seems to have affected more the manufacturing sector than the wide-economy, whereas the 1990 dip in the wage gap due to the *Collor Plan* was not so strong in the manufacturing sector.

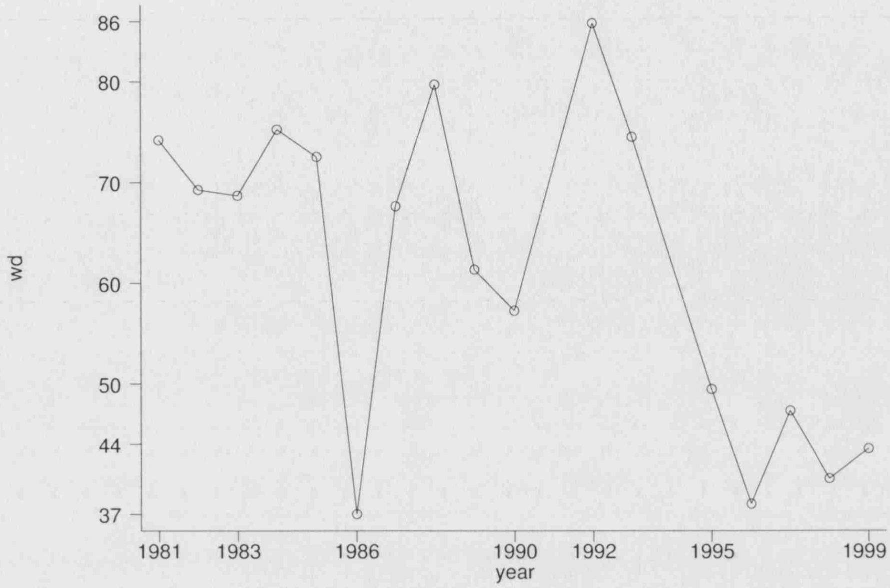

Figure 4.2: Wage Differential between Registered and Non-registered Workers in the Manufacturing Sector (in %) - 1981 - 1999

aim is to establish whether the coincidence between the reduction in the wage differential between registered and non-registered workers and the fall in the proportion of registered workers and the end of the trade reforms was a mere coincidence or was really linked to the trade liberalisation measures. In the next section, we discuss the impact of trade liberalisation in developing countries labour market, and how it might have affected the wage differential between registered and non-registered workers in Brazil and the proportion of informal sector workers.

4.3 Trade Liberalisation and Labour market

4.3.1 A Brief Overview of the Literature on the Impact of Trade Liberalisation on Developing Country Labour Markets.

Researchers have looked for evidence of any of the Hecksher-Ohlin/Stolper-Samuelson (HOS) framework's implications. More specifically, they have looked for evidence that trade liberalisation has triggered the following sequence of events: 1) increases in the relative price of less-skill intensive products, firms or industries; 2) a positive effect of these price increases on the demand for unskilled workers; 3) a reduction in the wage premium of skilled workers, leading to a lower wage inequality in these countries, and 4) an increase in the share of skilled workers in all industries. Such chain of causation relies on the hypothesis that developing countries have a higher endowment of unskilled workers, leading to comparative advantages in the production of goods intensive in this factor. The Hecksher-Ohlin theorem states that a country will tend to export goods that are relatively intensive in the abundant factor. The Stolper-Samuelson theorem shows that changes in the output price have a more than proportional effect on the return of the relatively abundant factor in the industry where the shock occurred. The combination of these two theorems yields the above prediction that trade policy changes that lead to a higher relative price of unskilled-intensive goods should bring about an increase in the relative wage of unskilled workers[11].

Assuming the special case where the functional form of the production functions for all sectors and for the aggregate utility function is Cobb-Douglas, the proportional change in the relative wage rate between skilled workers (s) and unskilled workers (u) in an open economy can be expressed as:

$$\widehat{\left(\frac{W_s}{W_u}\right)} = \frac{1}{\beta_1 - \beta_2} \widehat{\left(\frac{P_1 A_1}{P_2 A_2}\right)} \tag{4.1}$$

where β_1 and β_2 are the proportion of skilled workers in the skilled-intensive sector and in the unskilled-intensive sector, respectively; A_1 and A_2 are technology parameters in these same sectors, and P_1 and P_2 the respective product prices. Since $\beta_1 > \beta_2$ changes in prices and/or technology have a more than proportional effect on changes in the relative wage, an increase in P_2, the price of the product in the unskilled-intensive sector, should lead to a more than proportional fall in the relative wage of the skilled workers.

However the above result only holds if we assume that: 1) the economy is small so that it cannot affect the international price of the product, which is assumed to be exogenous, 2) the economy is inside the cone of diversification, meaning that tradable goods intensive in both factors, skilled and unskilled labour, are produced in that economy; 3) there is no product differentiation, i.e., foreign and domestic goods are perfect substitutes; and 4) there are no mobility barriers for workers to respond to wage changes. A corollary of this theory is that changes in the supply of different factors do not alter their relative prices (as changes in the relative price do). Changes in the factor endowment of a country would increase the production

[11]Due to this sort of reasoning trade liberalisation has been prescribed as a policy that not only would increase economic efficiency, but also would deliver less inequality in developing and unskilled labour abundant countries.

in the industries intensive in the factor, without altering its relative price (Johnson and Stafford, 1999).

So far the empirical literature on the developing countries has found at best mixed results regarding HOS predictions. For the Mexican experience in the mid-1980s, Hanson and Harrison (1999) show that the reduction in tariff protection disproportionately affected low-skilled industries, contrary to what one would expect for a developing country. However, they argue that such a result does not come as a surprise, since the Mexican import substitution strategy had extended trade protection preferentially to industries that made relatively intensive use of unskilled labour. The relative higher protection of industries in which the countries would, in principle, have comparative advantage was also noticed by Currie and Harrison (1997) for the Moroccan manufacturing sector. Goldberg and Pavcnik (2003) also found that the structure of tariff protection benefited more the industries with a higher share of unskilled workers in Colombia and in Brazil[12]. These studies highlight the necessity to understand the previous structure of protection before assuming that any trade liberalisation reform would trigger a reduction in wage inequality, as measured by the relative wage between skilled and unskilled workers. If the protected sectors were the ones in which the country already had comparative advantage, then the openness measures could lead to a fall in their product prices and then trigger an increase in the relative demand for the scarce factor (skilled workers).

Behrman et al. (2001) do not find evidence that trade liberalization has any overall widening effect on wage differentials for a panel of 18 Latin American countries - including Brazil - for the period 1977 to 1998. Robbins (1996a) also fails to find any relationship between trade liberalisation and wage inequality for Colombia. Gindling and Robbins (2001) find evidences consistent with a positive correlation between trade liberalisation and higher returns to education in Chile. Galiani and Sanguinetti (2001) find that manufacturing sectors where the import penetration increased the most, wage inequality also widened relatively more in favour of the most skilled workers in Argentina. However, they point out that the effect is quantitatively too small to explain the increase in the inequality in that country in the nineties. Pavcnik et al. (2002) show that the increase in the return of the college-educated workers coincides with the trade liberalisation in Brazil. They do not find any relationship between trade related measures and the increase in wage premium in sectors more affected by the reform, but they do find that the sector specific skill premium did rise for skilled workers. Green et al. (2001) also stress the coinci-

[12]Gonzaga et al. (2002a) found exactly the opposite result for Brazil. The reason for this difference, however, is not very clear.

dence between trade liberalisation in Brazil and the increase in the relative wage of college-educated workers, but fail to find any causal relationship. Unlike Pavcnik et al. (2002), however, they do find that the wage premium increased in sectors more affected by the trade reform[13]. Dickerson et al. (2001) using a pseudo panel approach find that the returns to education for college-educated workers fell after the trade liberalisation in Brazil, but do not find any correlation between trade measures and the return to education for college workers. Note that this result is at odds with Green et al. (2001). This is so due to the fact that the pseudo-cohort approach adopted by Dickerson et al. (2001) points to an overestimation of the returns to education yielded by the OLS method employed in Green et al. (2001). Arbache et al. (2004) reinforce the results in Green et al. (2001) and argue that within the traded sector, increasing openness was associated with lower wages but the downward impact of openness on wages was insignificant at the highest two education levels. Gonzaga et al (2002a) argue that the wage differential between skilled and non-skilled workers fell after trade liberalisation in Brazil and that the mechanism of transmission of this fall through tariff to prices and prices to wages is in line with HOS predictions[14].

The lack of strong evidence for HOS implications in developing countries has been rationalised via three hypotheses. The first is related to the perception that developing countries with higher proportion of semi-skilled workers may have been suffering strong competition from countries with a higher proportion of unskilled workers (Wood, 1997 and Hanson and Harrison, 1999). This halfway position of some industrialised developing countries, mainly in Latin America, would prevent HOS framework from working[15].

The second hypothesis assumes that trade may have caused a higher contact with leading-edge technology. In order to install this newly available technology, firms might have demanded more skilled workers to operate them and to adapt the production process to this more efficient technology. Such reasoning is advocated by the *skill-enhancing trade hypothesis* due to capital-skill complementarity (Robbins,

[13]In another paper Arbache and Corseuil (2003) find the opposite result using a fixed-effect model with time dummies. The result in Green et al. (2001) is based on the spearman correlation of long changes of the two variables.

[14]Notice that unlike Pavcnik et al. (2002) they argue that tariffs fell more in the most skill-intensive sectors, so that their analysis could follow the HOS mechanism.

[15]Davis (1996) argues that a possible explanation for this phenomenon is the fact that what is relevant to an economy is its relative skilled/unskilled labour supply within its cone of diversification. A country regarded as abundant in unskilled labour globally, may be in fact relatively skill abundant inside its relevant cone of diversification. In this case, trade liberalisation may have the effect of increasing inequality if the country starts suffering a stronger competition from relatively unskilled abundant countries within its cone of diversification.

1996b)[16] and by the *learning-by-trade hypothesis* (Pissarides, 1997). According to the latter, even if the technology is not skilled biased, in the short run the acquisition and incorporation of the technology would imply in a higher demand for skilled workers[17] The diffusion of more efficient technologies in developing countries would lead to an increase in A_1 in equation (4.1) and then to an increase in the relative wage of skilled workers[18].

The third hypothesis argues that empirical studies fail to find any HOS trade related impact on income distribution because of global/pervasive skill biased technological change (Berman and Machin, 2000). In this case, no correlation between trade measures and the increase in the premium of skilled workers would be observed, since this would be an economy-wide phenomenon.

A second point that has been emphasized in the literature is the role of trade liberalisation in changing institutional features of industrial relations and then indirectly affecting the wage and employment structure (mainly in the manufacturing sector) of countries that have undertaken trade reforms. The main hypothesis has to do with the loss of union power triggered by trade reforms. Both the fall of trade barriers and tariff reductions increase the price elasticity of product demand, hence reducing rents that sustained the union wage premium.

Nickell (1999) points out that the relationship between wages and factors that reflect market power, such as market share, the number of a firm's competitors, reductions in tariff barriers, deregulation or the price of competing imports, have been used to assess the hypothesis that product market power affects wages.

We will briefly report some results of empirical studies on the impact of trade reforms on relative wages. Driffil et al. (1998) show that when non-tariff barriers were reduced in Britain, the wages in the relevant establishments fell significantly. Revenga (1992) observes a similar effect associated with falls in import prices in the US. Lang (1998) argues that the small effect of trade liberalization in New Zealand on

[16] According to this view, sectors under pressure to increase their productivity due to an increase in foreign competition would invest in technology that would lead to a higher demand for skilled labor. Evidence that sectors most affected by the trade liberalisation were the ones that experienced an increase in their wage premium or where the wage premium of skilled workers increased most would be in line with this argument.

[17] This argument is also found in the theory of technological cycles, where the development and introduction of new technologies would generate a rise in the demand for skilled workers, but as the technology becomes standard and all workers learn how to work with it, the shock fades away. See Goldin and Katz (1998).

[18] Such explanation is in line with the simultaneous rise in inequality in both developed and developing countries. However, in order to observe all these effects it is necessary that the industry where the transference of technology occurred is still operating in the developed countries, so that the fall in prices (P_1) does not compensate for the increase in A_1. See Johnson and Stafford (1999) section 4.8 for a brief discussion of these models.

the composition of employment suggests that the effect of tariffs on wages and firms' monopoly power, reducing both of them, eliminated any effect on the distribution of employment. Borjas and Ramey (1995) show that the impact of international trade on relative wages (skilled/unskilled wages) depends on the market structure of the industry affected. They argue that many of durable good industries in the US in the 1980's that employed a disproportionate share of less educated workers were highly concentrated, earned significant rents, and shared those rents with their workers by paying them higher-than-average wages. Their empirical evidence shows that employment changes in a small group of trade-impacted concentrated industries can explain not only part of the aggregate rise in wage inequality in the US, but also some of the differences in trends in wage inequality across metropolitan areas. Somewhat against those findings, Johnson and Stafford (1999) in their review of the impact of trade on labour market institutions argued that despite the theoretical negative relationship between increased international competition and "monopoly rents" enjoyed by the firms protected in the past, there is no strong evidence of a negative effect of increased trade on unionism either in the US or in the UK.

The findings on the impact of trade-related variables on wages and on employment for developing countries also tend to place this sort of rent-sharing argument as a possible explanation for decreases in the average wage, at least, in the manufacturing sector. Arbache (1999) argues that the market-oriented reforms in Brazil, and particularly trade liberalisation, led to a higher demand for skilled workers that ended up increasing union power. This happened because unlike developed countries, the unionized workers are relatively more educated in Brazil than non-unionized workers. Revenga (1997) finds evidences that the (negative) impact of trade liberalisation on wages in Mexico was higher than the (negative) effect on employment. She argues that this fact may be explained by the prevalence of rent-sharing schemes in the period previous to trade liberalisation. Such schemes would have allowed unions and firms to agree in cutting "excessive wages" rather than adjusting the employment margin after the reforms. Similarly, Currie and Harrison (1997) analysing the Moroccan trade liberalisation argue that in an imperfect competitive framework where some rents were captured by workers in the form of higher wages, firms could also respond to their rent loss by cutting wages and substituting temporary workers for permanent ones. Menezes-Filho and Arbache (2002) show evidences of rent-sharing for unionized workers in the manufacturing sector in Brazil. However, they also find that the increase in quasi-rents brought about by trade liberalization was not shared with unionized workers[19]

[19]Note that the effect of trade liberalisation on rents or quasi-rents is not consensual. It can

4.3.2 Trade Liberalisation and Segmented Labour Market.

The impact of trade liberalisation on the informal sector is less understood and less documented. Behrman (1999) points out that most studies have focused on the impact of trade liberalisation on the formal manufacturing sector, but less is known about its effect on the informal sector, both in the manufacturing sector[20] and in the entire labour market. In this section, we will highlight possible effects of trade liberalisation on segmentation – in terms of the registered *versus* non-registered classification - in Brazil. First, we will make a parallel between the formal/informal wage gap and the skilled/unskilled wage gap prediction of the HOS. Then, we will relate it to the idea that the informal sector can be related to the non-tradable sector in Brazil, so that its dynamics is only determined by the supply and demand in that specific sector. Finally, we will relate formal/informal dichotomy to the ability of capture rents by more well organized/proctected workers. Our main argument is that trade liberalization may have reduced the rents that protected sectors would share with their workers, leading to a fall in the wage gap.

The first thing to notice about the relationship between trade liberalisation and the relative wage of registered and non-registered workers is that such a classification is an institutional feature and not a skill-based classification. Unlike the classifications by educational attainment; occupational categories; production and non-production workers as traditionally used in the literature and that are based on productivity related features; the classification between registered and non-registered workers parallels the classifications between union/non-union workers and/or temporary/permanent workers, which are much more related to institutional features of the labour market. Nevertheless, given the lower level of schooling of non-registered workers[21], it is common to to assume that the production function of the industries may be approximately depicted as having registered and non-registered workers as two inputs instead of the traditional division between skilled and unskilled workers. This argument can be made stronger if we assume that it is the employer's decision that determines the *institutional* status of the worker and employers will choose workers that, given their productivity, minimise their production costs. Thus assuming that the employers know better than the researcher the productivity of the workers, the employees *institutional status*, i.e, whether he/she is registered or not,

decrease rents because it increases competition, or it can increase them because it forces an increase in productivity. Thus, the overall effect depends on which of these two effects will be dominant.

[20]To the best of our knowledge the only paper to address this issue is Goldberg and Pavcnik (2003).

[21]As seen in Chapter 2, this difference is persistent in the data and the difference in means is statistically significant at conventional levels.

may well be determined by their productivity[22].

In this context, equation (4.1) might give us the prediction of the impact of price changes (ΔP_i) or technological changes (ΔA_i) on the relative wage of registered/non-registered workers. Assuming that both factors are used in the tradable industries, and that there is at least one industry where the registered status is prevalent, and another where the non-registered status is prevalent, an increase in the price of non-registered prevalent industry good would lower the wage differential between registered and non-registered workers. The problem with this reasoning, as already mentioned above, is that in contrast with the other common classifications, the non-registered status is not necessarily production-related. Factors such as lower school attainment do not necessarily mean that non-registered status may be directly linked to lower observed or non-observed skills. However, it is possible that on average what happens with the wage of unskilled workers is reflected on non-registered workers[23].

The presence of a non-tradable sector where non-registered status is prevalent would not affect such analysis, i.e., the world prices of good 1 and 2 would still determine the relative wage as in equation (4.1). However, the presence of a non-tradable sector with such characteristic would make the cone of diversification of that country thinner, and more likely that a fall in the relative price of the "non-registered" prevalent good would lead that industry to close down so that non-registered workers would be employed only in the non-tradable sector (Johnson and Stafford, 1999). In this case, equation (4.1) would no longer represent the skilled/unskilled (or registered/non-registered) wage rate. The relative wage would be determined as in the case of a closed economy:

$$\widehat{\left(\frac{W_s}{W_u}\right)} = \frac{\beta_4}{1-\beta_4}\left(\frac{U_4}{S_4}\right) \tag{4.2}$$

where S indexes skilled workers and U indexes unskilled workers, the subscript 4 stands for the non-tradable sector 4, which is "non-registered" prevalent and unskilled intensive.

It is clear from (4.2) that regardless of the changes in prices in the tradable sector, the relative wage would be unaffected. In this context, one should not expect to find any effect of trade-related variables on the relative wage of skilled/unskilled workers

[22] In Chapter 3 we tested the hypothesis of a job queue for formal jobs. The employers decision of hiring a worker as registered was modelled as well as the employee's decision of joining the queue. The hypothesis of the non-existence of a job queue for registered jobs was strongly rejected.

[23] In fact, the literature using Brazilian data has failed to find strong evidence of lowering wage inequality or a fall in the wage premium of skilled workers after trade liberalisation. See for instance Green et al (2001) and Dickerson et al. (2001). The only empirical paper that claims such effect is Gongaza et al. (2002a). See last subsection.

or registered/non-registered workers. Midway situations between the result of the closed model represented by equation (4.2) and the open model in equation (4.1) arise if one assumes that a) domestic and foreign goods are not perfectly substitutes and b) labour types cannot move in response to wage changes (Johnson and Stafford, 1999).

In fact, some commentators have used the tradable *versus* non-tradable approach to explain the dichotomy between formal *versus* informal sector in Brazil[24]. The idea behind this correspondence is that the earnings of workers in the informal (non-tradable sector) is determined by supply and demand in that sector, whereas the earnings of workers in the formal sector is determined by the external demand for the export goods. The demand for the national manufactured product would be lower after the trade reform due to the access to cheaper products, whereas the non-tradable sector would be protected from that competition. The difficulty in such argument is how to justify the lack of mobility between workers from the formal to the informal sector. It could be argued that the informal sector would act as a cushion for workers displaced from the tradable sector, leading to a downward pressure on wages in that sector. Therefore, it is not clear how trade liberalisation would affect the wage of the employees in the tradable sector, but would not affect the wage of the employees in the non-tradable sector[25].

This sort of argument would be more justifiable in a context of imperfect competition where one would focus on the effect of trade reforms on institutional features of developing country labour markets. In this framework, trade reform would squeeze rents that would have been captured by protected firms and shared with their employees. As seen above, if the protected firms were the ones abundant in the scarce factor (skilled labour) then one should expect under HOS assumptions that the reduction or elimination of trade barriers would reallocate resources to the now more competitive firms based on the abundant factor and hence increase the demand for unskilled labour. However, as many papers have shown this is not necessarily the case, and the protected sector may have been in fact the one that had a higher proportion of the abundant factor. Alternatively, if the protected sector cannot be characterised as perfectly competitive, then there is space for some sort of rent sharing. Therefore, workers in most affected industries would experience a reduction in their bargaining power –and so in their wages - since the ground for rent sharing

[24]See Barros et al. (1998) for an informal explanation of the argument.

[25]Generally, it is argued that the participation of non-registered workers in the manufacturing sector is so small that it can be considered as a specific factor of the non-tradable sector. However, as seen in the last section, despite the higher proportion of registered workers in the manufacturing (tradable) sector, its proportion has fallen over time.

would be reduced. Assuming that the most protected sectors are the ones with a higher proportion of registered workers, or where registered workers profit more from rents thanks to the market power of their firms, one should observe a reduction in the wage premium for registered workers after the reforms.

In Brazil, as seen in Chapter 2, registered workers are more likely to be unionized and to work in large firms, which are more likely to have market power and therefore to have some loss due to trade liberalisation. Therefore, it would be reasonable to assume that the degree of segmentation, as measured by the wage premium for registered workers, is positively correlated with trade protection measures[26]. The process of trade liberalisation should hence curb the wage differential between registered and non-registered workers. Besides this direct effect, skilled workers displaced from registered jobs may have joined the pool of non-registered workers, increasing their average skill level and contributing to the reduction in the wage differential due to changes in the composition of the two groups[27] However, it is also important to consider that the effect of trade liberalisation on productivity[28] may, on the other hand, lead to higher wages for the workers who managed to keep their jobs and also to an expansion of registered jobs. This process could lead both to a higher proportion of registered workers and a larger wage gap between the two groups. Therefore, the impact of the trade liberalisation on the wage gap is mainly an empirical question.

So far we have focused on the wage differential, but as mentioned above, trade liberalisation can also affect the allocation of different types of workers between and within industries. In particular, Brazilian firms may have reacted to the trade liberalisation shock not only by substituting non-registered workers for registered workers, but also sub-contracting part of the tasks that they could have performed

[26]See the discussion about the correlation of wage premium and trade-related measures in the next section.

[27]Barros et al. (1998) using the Oaxaca-Blinder decomposition show that composition effect (change in the attributes) was responsible for 25% of the reduction in the wage differential and conclude that the increase in the return of non-registered workers productive characteristics, possibly due to a positive demand shock, was the main factor responsible for the reduction in the wage gap. Similar results are reported in Chapter 2 of this thesis.

[28]Muendler (2001) using an unbalanced panel of medium and large manufacturing firms in Brazil estimates the impact of nominal tariffs and market penetration on total factor productivity at firm-level. He instruments nominal tariff and market penetration with inflation rate and real exchange rate and finds that lowering tariffs by 10% leads to an increase of almost 0.3% in productivity, while a 10% point increase in market penetration leads to another 1.3% increase in productivity. It is also worth noting that the changes in the probability of transition between categories of activity was, according to Muendler's decomposition exercise, responsible for 50% of the increase in productivity observed in the period. Hay (1999) and Rossi Junior and Ferreira (1999) also find a positive relationship between manufacturing productivity and trade liberalisation in Brazil.

earlier in an attempt to reduce costs[29]. The reallocation of part of the production to smaller firms may have led to a higher participation of the non-registered workers in the pool of manufacturing employee. Similarly, the reduction in absolute terms of the number of jobs in the manufacturing sector due to adjustments in the size and composition of its workforce may have led to an increase in the proportion of non-registered workers in the entire economy.

Most of these channels are hard to assess since there is no available and compatible data on "informal firms" in Brazil that would allow one to compare its performance over the recent period with the performance of medium to large firms[30]. Nevertheless, data from the Brazilian annual household survey (PNAD) shows that the proportion of non-registered workers increased in the manufacturing sector from 15% to 24% from 1981 to 1999, which is a clear indication of a possible lower degree of compliance within the manufacturing sector. Furthermore the proportion of workers in small firms (up to ten employees) increased from 40% to 50% in the entire economy and from 15% to 23% in the manufacturing sector during the same period. These changes may have two causes: a) more firms decided to contract workers illegally; b) the balance on the birth and death of firms favoured smaller firms that are more likely to employ non-registered workers. Somewhat supporting this latter hypothesis, Muendler (2001), based on an unbalanced panel of medium to large size manufacturing firms, finds that in the period 1992-1998, the probability of transition from active status to extinct (shut down), and from suspended to extinct had increased considerably in comparison to the period 1986-1990. This evidence lends some support to the argument that there was a clean up effect among medium to large firms after trade liberalisation that may have led to a reduction in the proportion of registered workers. In the next section we will describe the main characteristics of the trade liberalisation process.

4.4 Some Features of the Trade Liberalisation Process in Brazil

The process of trade liberalisation in Brazil started in 1988, but only gained pace after 1990. The structure of protection of the national production was very complex. Besides high tariffs, it included several non-tariff barriers along with special

[29] Maloney (1997) shows a sharp increase in the number of workers who earn by piece, the so-called contract workers, after the trade liberalisation in Mexico.

[30] Most of the available data on firms such as value-added, payroll payment, working hours are data collect from medium and large firms.

regimes that exempted some sectors of paying standard import tax. One of the most important restrictions was the "law of national similar" which forbade the import of any product that had an equivalent/similar produced in Brazil (Moreira and Correia, 1998). There were also import quotas for firms and the need to get permission one-year in advance to import.

According to Kume et al. (2000) the trade liberalisation process in Brazil can be divided in three phases. The first period between 1988-1989 consisted in scrapping redundancies in tariffs, i.e., in cutting the excessive and unnecessary level of tariffs that is more than sufficient to compensate for the difference between the world price and the domestic price, and in the partial reduction of special regimes[31]. The second period 1990-1993 witnessed the elimination of all special regimes for imports except for *Zona Franca de Manaus* (Manaus Duty Free Zone) and for the computer industry; the elimination of all non-tariff barriers and a schedule for a gradual reduction of tariffs during four years from 1990 onwards. The programme intended to reduce by a lower amount the tariffs of sectors more intensive in technology such as computer and chemical industries and those with a high demand for national input products such as the automobile industry. It also intended to reduce more sharply, in the beginning of the reform, the tariffs on capital and intermediate goods, and later on the tariffs on consumption goods. However, in order to tighten up the price control of oligopoly groups and curb the inflationary process, the government decided to anticipate the lowering of tariffs for consumption goods. For this reason, in October 1992, the tariffs were reduced to the levels that were scheduled for January 1993 and in July 1993 to the levels that were scheduled for January 1994. Such anticipations announced in February 1992 reinforced the natural experiment environment in which the process of trade liberalisation took place in Brazil. The third and last period consisted of the anticipation of the Mercosur Common Tariff from January 1995 to September 1994 and of the reduction of tariffs to 0% for intermediate goods and 2% for consumption goods which figured prominently in the bundle of goods that compounded the price indices in order to control inflation just after the *Plano Real*[32].

[31]According to Kume et al. (2000) these partial reforms did not have a strong impact on the foreign competition due to their small scope. Despite the apparent strong reduction in nominal tariffs, most of them were redundant. In addition, the most important non-tariff barriers were not lifted during this period.

[32]The appreciation of the *Real* witnessed after the *Plano Real* together with reductions in tariffs and an increase in the aggregate demand led to successive trade deficits. This fact led the government to slightly increase tariffs in 1995. It is worth noting that whereas the tariffs were used to control inflation rather than to enhance productivity in the first phases of the process, after the *Plano Real* it was used as a mean to smooth the increasing trade deficits faced by the country. However, due to the agreements established by the Mercosur, the ability of the government to

	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Average tariff	54.9	37.7	29.4	27.2	20.9	14.1	12.5	10.2	10.8	10.8	13.4	13.4
Standard deviation	21.3	14.6	15.8	14.9	12.7	8.2	6.7	5.9	7.4	8.7	7.6	6.6
Maximum tariff	102.7	76	75	78.7	58.7	39	34	23.5	41	52.4	47.1	38.1
Minimum tariff	15.6	5.6	1.9	3.3	1.7	0.6	0.0	0.0	0.0	0.0	0.0	0.0
Import Penetration			4.4	4.2	6.0	6.1	6.9	7.7	10.0	10.7	12.5	13

Sources: Kume et al. (2000) for the average Nominal Tariff and Fonseca et al. (2000) for the import penetration coefficient in the manufacturing sector.

Table 4.1: Nominal Tariff and Import Penetration Coefficient (in %)

The partial reversion in the tariff reduction observed from 1995 onwards was followed by other measures that intended to reduce the trade deficit. Among these measures were the reintroduction of some non-tariff barriers such as previous authorisation for some imports and increases and a slight increase in tariff as can be observed in Table 4.1. However, in spite of a short-term reduction in the growth rate of the import penetration ratio, these measures were not able to prevent the increase in the foreign market penetration.

Table 4.1 reports the evolution of the nominal tariff between 1987 and 1998 and of the import penetration coefficient between 1989 and 1998[33]. As one can see, not only did the process of trade liberalisation lead to a decrease in the average value of nominal tariffs, but also to a reduction in the dispersion of their level among industries. In 1987, the average nominal tariff was 55% with a standard deviation of 21.3. These figures were down to 13.4% and 6.6, respectively in 1998.

Despite this sharp fall in tariffs, imports did not increase immediately. This occurred because the process of trade liberalisation took place during the recession period 1990-92 and in a period when the currency was depreciated[34]. Furthermore, as mentioned above, a considerable part of the reduction in the late 1980's was due to cuts in redundant tariffs. The import participation in the total supply increased from 3.78% in 1985 to 4.05% in 1992, and by 1997, this figure was up to 5.67% (Oliveira-Junior, 2000). Similarly, the import penetration ratio was 6.1% in 1992 (the same level as in 1985) and increased to 13% by 1998 (Fonseca et al., 2000).

An important point to note is the possible lack of a direct relationship between either the nominal tariff or effective tariff[35] and import penetration ratio across

change tariffs was constrained.

[33]All the data on the trade measures are in Appendix B for all years. The data in Kume et al. (2002) was made compatible with the PNAD industry classification by averaging sectors using as weight their added-value.

[34]In our empirical strategy we tackle the possible lag between tariffs and import penetration using lags for tariffs when it enters the regression as an explanatory variable and through controls for business cycle (time dummies) and specific exchange rates

[35]Effective tariff is a measure that not only includes the nominal tariff for the output of the industry, but also takes into account input tariffs as a way to measure the overall degree of protection enjoyed by the industry. See Kume et al. (2000) for more details.

	Nominal Tariff	Import Penetration	Export Orientation	Wage Premium	Proportion Registered
Effective tariff	0.929* *(0.000)*	-0.042 *(0.874)*	0.324 *(0.205)*	-0.005 *(0.985)*	-0.049 *(0.852)*
Nominal tariff		-0.164 *(0.529)*	0.172 *(0.510)*	-0.206 *(0.428)*	-0.213 *(0.411)*
Import penetration			0.328 *(0.198)*	0.238 *(0.358)*	0.444*** *(0.075)*
Export Orientation				0.267 *(0.299)*	0.086 *(0.743)*
Wage premium					0.064 *(0.808)*

*Significant at 1%; **Significant at 5%; ***Significant at 10%
P-value in parenthesis

Table 4.2: Spearman Rank Cross-correlation in 1987

sectors. As discussed above, there were many non-tariff barriers preventing imports as well as special regimes that allowed the protected sector to import inputs without paying the due amount of import tax. Moreover the import penetration ratio captures other features that are not captured by tariff measures such as the level of the real exchange rate and tastes and preferences of consumers. Tables 4.2 and 4.3 show the Spearman rank correlation between all pairs of variables that we are going to investigate for 1987 and 1998, respectively. The variables are a) trade-related variables: nominal tariff, effective tariff, import penetration ratio[36], export orientation; and b) the wage premium for registered workers and the proportion of registered workers for 17 manufacturing industries[37]. It is striking the lack of rank correlation between import penetration ratio and effective tariff. In 1987 the rank correlation was slightly negative (-0.04), but not statistically significant; in 1998, the correlation was positive (0.22), and again not significant.

It is worth noting that the only statistically significant rank correlation for both years, besides the obvious correlation between effective and nominal tariff, was the positive correlation between the import penetration ratio and the proportion of registered workers. This surprising result indicates that manufacturing sectors that tend to have a higher import penetration ratio also have a higher proportion of registered workers.

[36]Import penetration ratio is measured as the ratio between imports over the sum of the value of the production plus imports minus exports (M/(Y+M-X)), which gives us a measure of the real external competitive environment faced by each industry in the internal market. This variable is important because it is determined not only by tariffs, but also by other policy instruments and macroeconomic variables such as non-tariff barriers, exchange rate and GDP growth that determine the actual degree of openness of the economy. Data for import penetration ratio and export orientation (exports/value of the production) come from Haguenauer et al. (1998).

[37]The manufacturing industry groups at 2-digit level are non-metallic, metallurgic, mechanic, electronics, transport, wood and furniture, paper printing and publishing, rubber, chemical, oil refining, pharmaceutical and perfume, plastics, textiles, clothes, shoes and leather, food, and other industries.

	Nominal Tariff	Import Penetration	Export Orientation	Wage Premium	Prop. Registered
Effective tariff	0.919 *(0.000)**	0.221 *(0.395)*	0.326 *(0.202)*	0.397 *(0.115)*	0.091 *(0.729)*
Nominal tariff		0.382 *(0.130)*	0.353 *(0.165)*	0.279 *(0.277)*	0.103 *(0.694)*
Import penetration			0.390 *(0.122)*	0.142 *(0.586)*	0.637*** *(0.006)*
Export. Orientation				0.257 *(0.319)*	0.088 *(0.736)*
Wage premium					-0.140 *(0.593)*

*Significant at 1%; **Significant at 5%; ***Significant at 10%
P-value in parenthesis

Table 4.3: Spearman Rank Cross-correlation in 1998

Another interesting feature is the rank correlation among these variables between 1987 and 1998. Table 4.4 shows that both effective and nominal tariffs have a low positive own-rank correlation between 1987 and 1998[38], the correlation for the effective tariff is not statistically significant and for the nominal tariff is only significant at 10% level. Similarly, the own-rank correlation for the wage premium is also low, despite being positive (0.45) and significant at 10%. In contrast, import penetration ratio, export orientation and the proportion of registered workers have a much higher own-rank correlation, 0.86, 0.90, 0.81, respectively, which are all significant at 1% level. Thus it seems that the changes in tariffs were strong enough to change the relative position of the 17 manufacturing industries in relation to effective and nominal tariffs and in relation to the wage premium. However, the same is not true for the export orientation, for the import penetration ratio and for the proportion of registered workers. Nevertheless, it is important to bear in mind that the rank-correlation does not take into account the intensity with which the different industries were affected, just their relative position within a specific year rank.

Figures 4.3 and 4.4 plot all the data points of our ten-year (1987, 1988, 1989, 1990, 1992, 1993, 1995, 1996, 1997 and 1998) and 17 industries panel for the relationship between wage premium and trade-related measures and for the relationship between proportion of registered workers and trade-related measures, respectively. Both export orientation and effective tariff are positively correlated with the wage premium, whereas there seems to be a negative correlation between import penetration ratio and wage premium[39] for registered workers as one would expect. As for the proportion of registered workers, surprisingly both effective tariff and import

[38]The own-rank correlations correspond to the numbers in the diagonal of Table 4.4.

[39]However, none of these correlations are statistically significant at 5% of significance, but the coefficient on effective tariff is significant at 10% level.

	Effective Tariff(98)	Nominal Tariff(98)	Import Penetration (98)	Export Orientation (98)	Wage Premium (98)	Prop. Registered (98)
Effective tariff (87)	0.353 *(0.165)*	0.373 *(0.141)*	0.047 *(0.859)*	0.270 *(0.295)*	-0.017 *(0.948)*	-0.164 *(0.529)*
Nominal tariff (87)	0.363 *(0.152)*	0.434*** *(0.082)*	-0.042 *(0.874)*	0.154 *(0.554)*	-0.243 *(0.348)*	-0.194 *(0.457)*
Import penetration (87)	-0.064 *(0.808)*	0.066 *(0.801)*	0.858* *(0.000)*	0.441*** *(0.076)*	0.588** *(0.013)*	0.052 *(0.845)*
Export. Orientation (87)	0.360 *(0.155)*	0.343 *(0.178)*	0.282 *(0.273)*	0.897* *(0.000)*	-0.007 *(0.978)*	0.409 *(0.103)*
Wage premium (87)	0.206 *(0.428)*	-0.010 *(0.970)*	0.140 *(0.593)*	0.309 *(0.228)*	0.446*** *(0.073)*	0.326 *(0.202)*
Prop. Registered (87)	0.199 *(0.445)*	0.262 *(0.3092)*	0.662** *(0.004)*	0.186 *(0.474)*	-0.064 *(0.808)*	0.814* *(0.0001)*

*Significant at 1%; **Significant at 5%; ***Significant at 10%
P-value in parenthesis

Table 4.4: Spearman Rank Correlations between 1987 and 1998

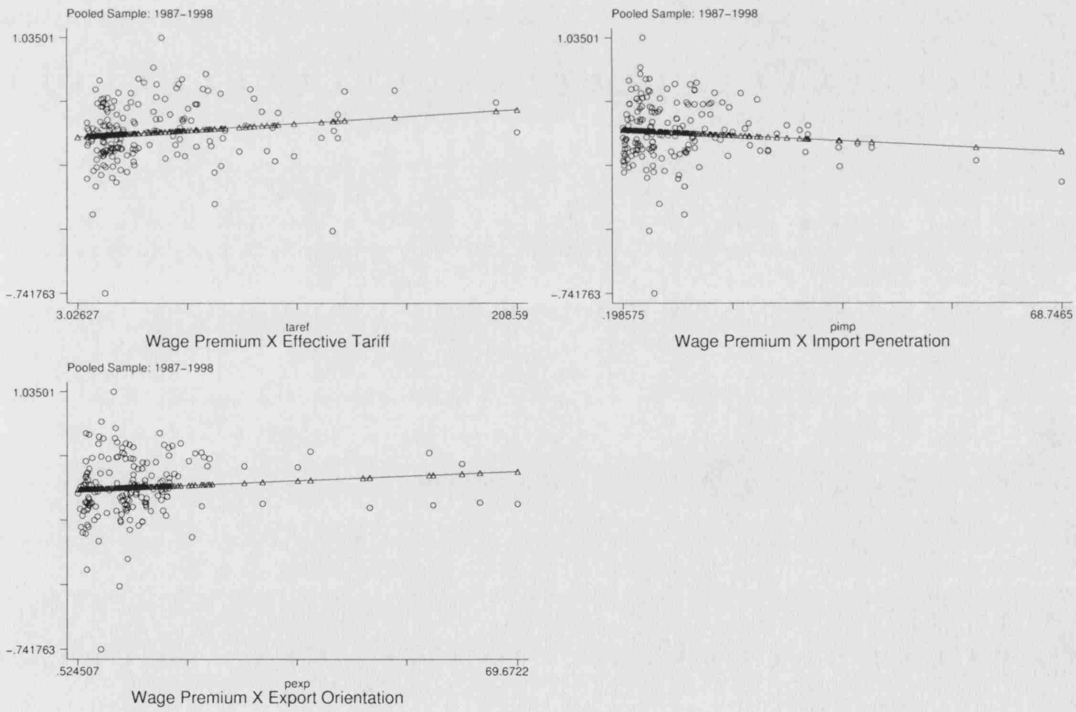

Figure 4.3: Wage Premium and Trade Measures (Pooled Sample) - 1981 - 1999

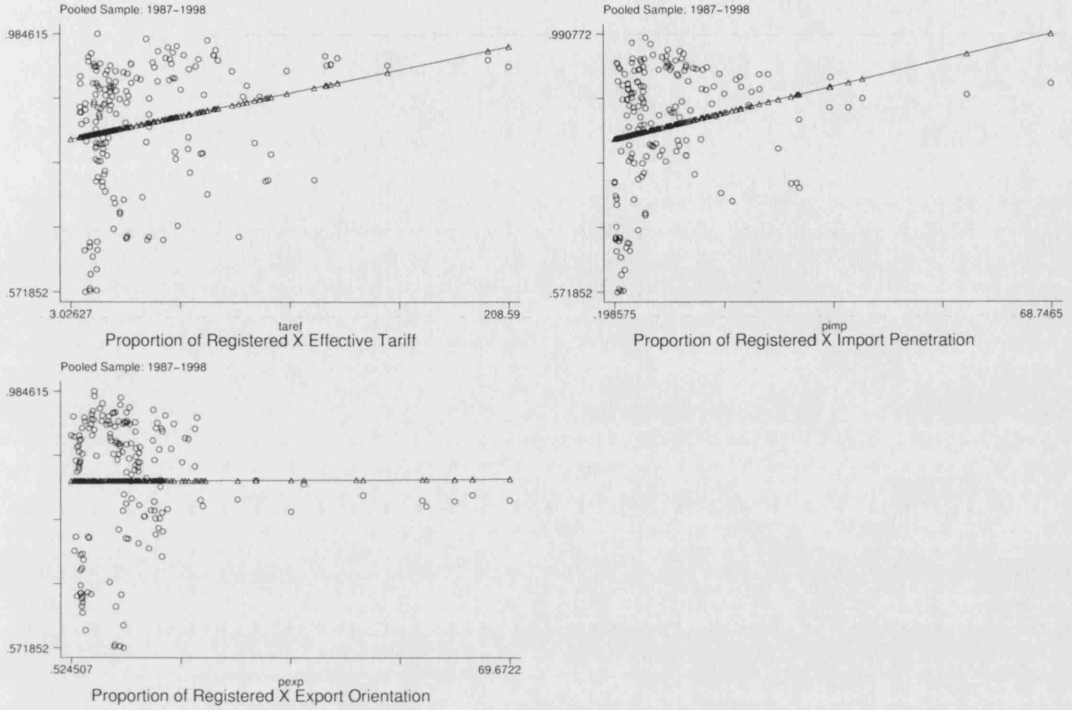

Figure 4.4: Proportion of Registered Workers and Trade Measures (Pooled Sample) - 1981 - 1999

penetration ratio show a strong positive correlation with it[40], whereas the export orientation ratio seems to be only slightly positively correlated with it, displaying an almost flat regression line. However, these results are just raw correlations and do not control for several other variables that are meant to affect both the wage premium and the proportion of registered workers. Moreover they do not control for industry fixed-effects and common macroeconomic shocks nor use weights in order to get more accurate correlations. Nevertheless, it anticipates the results we would expect to find from the regression analysis based on the pooled cross-section and time-series sample. In the next section, we will discuss the empirical strategies to investigate the relationship between trade-related variables and both the wage premium between registered and non-registered workers and the proportion of registered workers controlling for the drawbacks highlighted above.

[40]Both correlations are also statistically significant at 1% level.

4.5 Empirical Strategies

In order to test whether or not trade liberalisation did have an effect on the *degree of segmentation* and on the expansion of the informal sector in either the manufacturing sector or in the entire labour market, we apply three distinct reduced-form strategies. Since tariffs are uniform within the country in a given period of time, the first strategy relies on the variation of the degree of protection enjoyed by different industries, and on the different speed of the reform for different industries as described in the last section. This strategy allows us to check whether the industries most affected by the trade reform were also those that experienced the strongest reduction in both the *degree of segmentation* as measured by the wage differential and in the proportion of registered workers. Thus, in order to avoid bias due to the correlation between unobserved industry specific characteristics and trade related variables, we estimate this relationship using industry fixed-effect models and time dummies[41]. Controlling for time invariant unobserved industry characteristics is important because industry features that affect the relative wage of registered workers (and its proportion) may also affect their ability to lobby the government and/or the government priority in tariff reduction[42]. Likewise, time dummies would control for common macroeconomic shocks that would affect both the relative wage and the proportion of registered workers and the behaviour of trade related variables. For instance, during a recession it is likely that the import penetration ratio would fall as well as the relative wage of informal sector workers, whereas the proportion of workers in that sector increases. If the recessive period coincides with the trade liberalisation measures and we do not control for this common macroeconomic shock we would find a spurious relationship between the wage gap and the size of the informal sector and the trade measures.

The second strategy is based on the fact that the industries are not evenly distributed within the country. Thus if we assume that regions are differently affected by the trade reform according to their "industry composition" and that there is some sort of regional segmentation that prevents workers from migrating in the short term, we can test the existence of spillover effect of the trade liberalisation

[41]Unfortunately, due to the lack of the data for 1991 and 1994 for the variables related to wage differential and proportion of registered workers we cannot estimate first-difference models in order to test the robustness of the fixed-effect models.

[42]As mentioned in the last section, not only did the government establish different priorities in the tariff reduction schedule according to "observable" industry characteristics, such as the more aggressive reduction for intermediate and capital goods, but it also tried to curb inflationary pressures through reduction in tariffs due to the ability of certain industries to link prices to past inflation and wage hikes.

into the entire economy due to the different intensities with which each region was affected by the reform.

The third strategy adds two other sources of variation to estimate the impact of trade liberalisation on the proportion of registered workers and on their wage premium in the wide economy. We build industry-cohort cells that account for 30 sectors in the economy[43] - including both tradable and non-tradable sectors - and seven 10-year length cohorts. Thus we are able to exploit both cohort[44] and industry variation in order to identify the impact of trade liberalisation measures on employment structure and on the wage differential in the manufacturing sector. This is possible, because besides the different impact of the trade liberalisation on the manufacturing industries, we will be also using the non-tradable industries as a comparison group. These three identification strategies are detailed below.

4.5.1 The Impact of Trade Liberalisation on the Manufacturing Sector.

To test whether or not trade liberalisation had any impact on the fall in the wage differential between the registered and non-registered workers and on the increase in the proportion of the non-registered workers in the pool of manufacturing employees between 1987 and 1998, we run a fixed-effect model for both the coefficient of the "registered worker" dummy variable obtained from a standard semi-log Mincerian wage equation and the proportion of registered workers on a set of variables related to trade: effective tariff, nominal tariff[45], import penetration ratio and export orientation[46] for a panel of 17 industries[47]. The industry classification was developed in order to make the data from the household survey compatible with the trade-related data used in this paper. The sample is restricted to employed individuals between 14 and 65 years old with positive earnings and who worked more than 20 hours per week[48].

[43]This classification is not exhaustive of all sectors in the economy, because we excluded the public administration. We did that because public administration industry classification was used to filter out public servants that are classified as non-registered workers in the household survey. See Chapter 2 for a description of this filter.

[44]The use of cohorts enables us to control for industry-cohort fixed-effect, something that we could not do using only the repeated cross sections without building the cohorts.

[45]Data for effective and nominal tariffs come from Kume et al. (2000).

[46]Data for import penetration ratio and export orientation come from Haguenauer et al. (1998).

[47]The data for all these industries is available in the Appendix.

[48]We use the same cut off point employed in Chapter 2 to avoid the use of part time workers. However, we are aware that non-registered workers are over-represented among workers who work less than 20 hours, but overall workers working less than 20 hours represented on average less that 3% of the whole sample.

In the case of the effect of trade openness measures on the wage differential, we first estimate the following log wage equation for each pair of industry j and year t separately between 1987 and 1998.

$$w_{ijt} = \alpha_{jt} + \Gamma_{jt} X_{ijt} + \beta_{jt} Reg_{ijt} + \epsilon_{ijt} \tag{4.3}$$

where w_{ijt} is the log of the real hourly-wage for the individual i in industry j and year t and Γ is a vector of coefficients of the following independent variables X: region, gender, education (6 groups), experience, experience squared and metropolitan area, and β_{jt} is the coefficient for the dummy variable *Reg* that indicates whether the individual is a registered worker.

In a second step, we regress the estimate coefficient β_{jt} on the trade-related variables[49]:

$$\beta_{jt} = \alpha + \delta TM_{jt} + \phi_j + \theta_t + \varepsilon_{jt} \tag{4.4}$$

where TM_{jt} stands for trade measure variables in period j and time t and ϕ_t are industry dummies and θ_t are time period dummies and ε_{jt} is assumed to be a white noise.

The second reduced-form estimates refers to the impact of trade-related measures on the proportion of registered workers in the industries[50]:

$$Preg_{jt} = \alpha_2 + \lambda TM_{jt} + \phi_j + \theta_t + \varepsilon_{jt} \tag{4.5}$$

where $Preg_{jt}$ is the proportion of registered worker in industry j and year t and the remaining variables are as stated in equation (4.4).

Before discussing the results for this procedure, we will have a look at the long change (between 1987 – 1998) relationship between wage premium and proportion of registered workers and our trade-related measures. Figure 4.5 shows the long-change relationship in percentage between nominal tariff and the other three trade related measures used in this chapter. Nominal and effective tariff decreased in al-

[49]This two-step procedure has been used in the literature on the impact of trade on the US labour market by Borjas and Ramey (1994) and Slaughter (2001), see subsection 4.3.1 for a discussion of Borjas and Ramey (1994) results.

[50]One alternative to this approach would be to follow Goldberg and Pavcnik (2003) and regress the probability of being a registered worker against other trade measures. This implies a first stage where one would estimate the registered status against several controls such as schooling, experience and, most importantly, industry affiliation. In the second stage, the estimated coefficient of the industry affiliation would be the dependent variable in a regression equation like the one below, and the inverse of its sampling variance would be used as weight in the estimation. We checked the robustness of our results and found very similar results when applying this methodology.

most equal proportion for all industries. Export-orientation increased more in the industries whose nominal tariff decreased less, but the relationship is not statistically significant and in fact the regression line is almost flat. It is also worth noting that whereas the import penetration ratio increased for all industries, particularly for textiles and clothes, and the effective tariff fell for all sectors, the export orientation showed a less balanced figure, some sectors increased their export orientation and other experienced a reduction between 1987 and 1998. Figure 4.5 reveals a negative, but not statistically significant, relationship between changes in tariff and changes in import penetration ratio[51]. As mentioned in the last section, the strong reduction in nominal tariffs observed in 1988-89 was concentrated on redundant tariffs, this fact can explain the weak correlation between changes in nominal tariffs and changes in the import penetration ratio. Unfortunately, we do not have good data for several years on the non-tariff barriers by industry which were lifted in 1990. Therefore, we assume that changes in the import penetration ratio, despite its possible endogeneity[52] in relation to wage differential and to the proportion of registered workers, can act as a summary variable that captures both the effect of non-redundant tariff reductions and the elimination of non-tariff barriers.

Figure 4.6 shows the long changes between wage premium for registered workers and trade related measures in percentages. Changes in effective tariff as well as changes in the import penetration seem to be positively correlated with changes in wage premium ratio[53]. However, only the correlation between changes in effective tariff is statistically significant at 10%. Thus sectors that experienced the highest (long) fall in effective tariff were also the ones that witnessed the highest (long) fall in the wage premium. Export orientation is negatively correlated with changes in the wage premium, but it is not statistically significant. Changes in import penetration ratio are positively correlated with changes in the wage premium, which is again an unexpected result, since we would expected to find a negative impact of this openness measure on the wage differential between formal and informal workers.

Figure 4.7 shows a negative relationship between changes in effective tariff and import penetration ratio and changes in the proportion of registered workers. The export orientation seems to display a positive correlation, but the regression line is almost flat. Not surprisingly, none of these estimated correlations are statistically

[51]It is worth noting that this negative relationship is somewhat driven by two outliers: clothes and textiles industry (numbers 17 and 18, respectively, in the graphs).

[52]As a way to control for this problem in our regression analysis, we also try specifications using lags of the trade measures rather than their contemporaneous values.

[53]Notice that one should expected that the relationship between changes in import penetration ratio and changes in the wage premium would have the opposite sign of the relationship between changes in effective tariff and changes in wage premium.

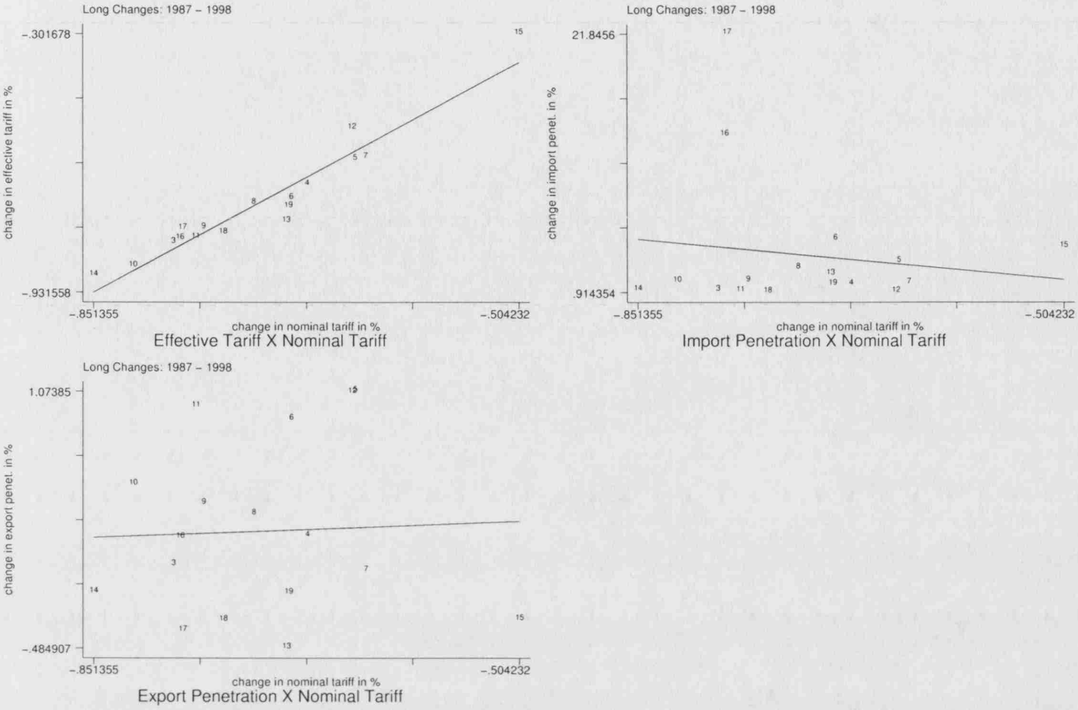

Figure 4.5: Nominal Tariff and Trade Measures (Long Changes: 1987-1998)

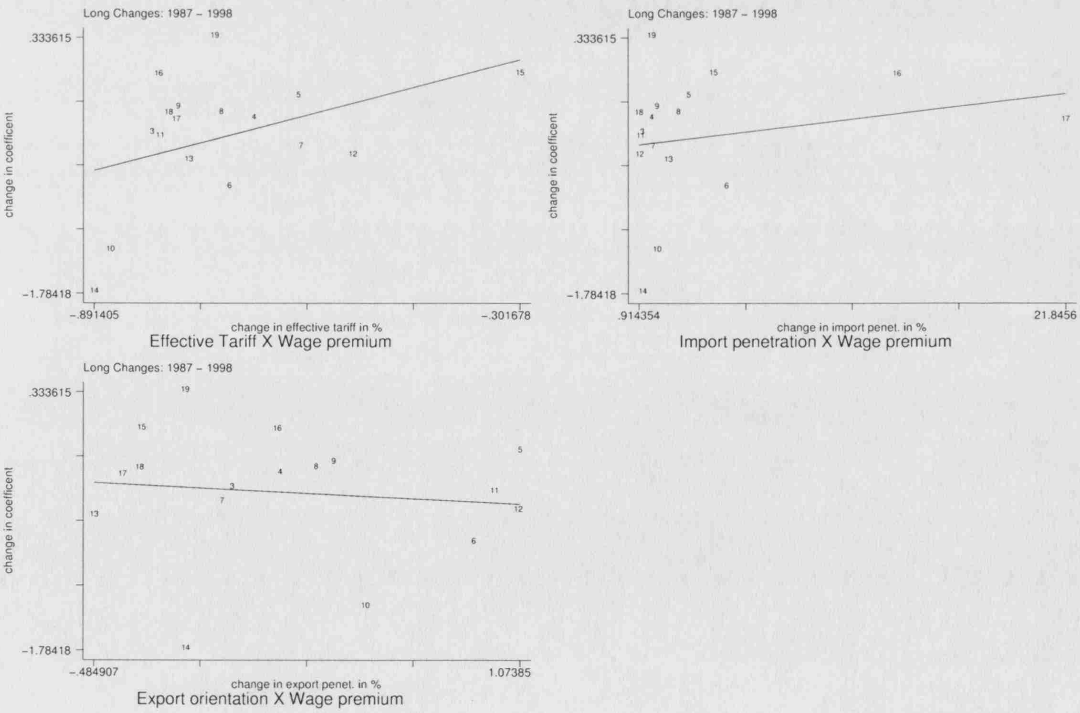

Figure 4.6: Wage Premium and Trade Measures (Long Changes: 1987-1998)

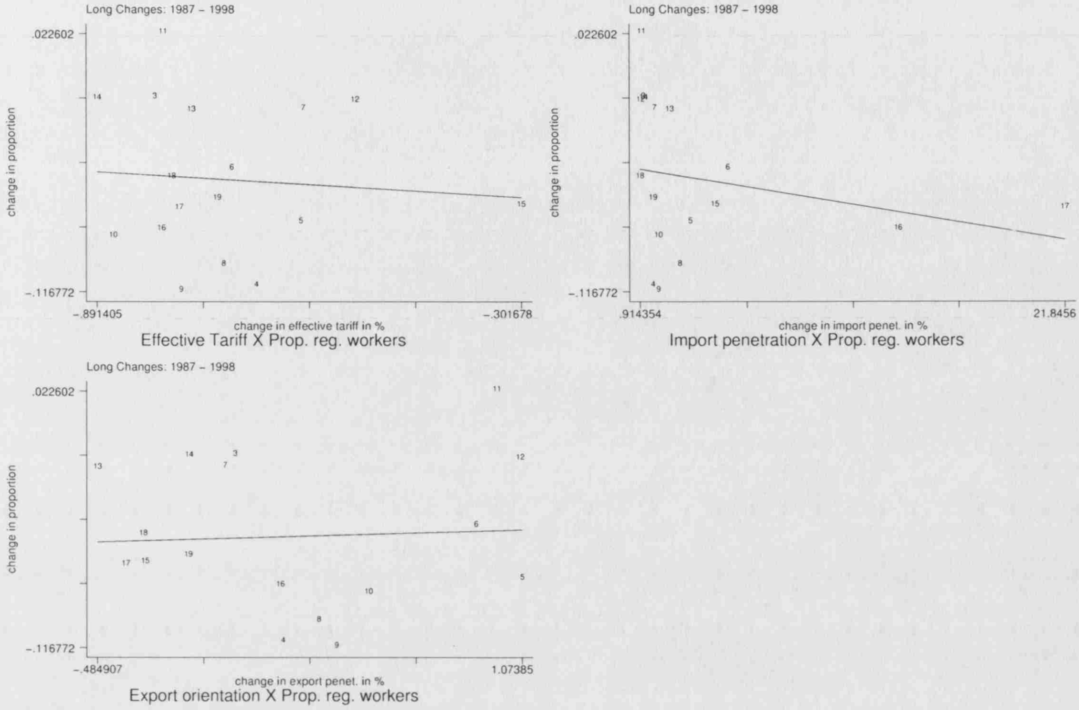

Figure 4.7: Proportion of Registered Workers and Trade Measures (Long Changes: 1987-1998)

significant. Therefore, the only statistically significant relationship found using the long-change approach was the negative impact that lowering tariffs had on the wage differential between registered and non-registered workers. However, despite controlling for individual fixed-effects, the long-change approach loses lots of variation in the estimation procedure, and for this reason we will emphasize the results of the fixed-effect estimations described bellow.

The fixed-effect regression versions of equation (4.4) are weighted by the inverse of the sampling variance of the dependent variable. Saxonhouse (1976) shows that estimations where the dependent variable is estimated in a first step and then regressed against a set of variables intended to explain it, suffer from heteroscedasticity because the stochastic term in the first stage is individual specific (in our case, industry specific). Similarly, equation (4.5) is estimated using the share of the industry in the total manufacturing employment as weight. Besides this weighting scheme, the standard errors in both equations are Huber-White corrected for any other source of general heteroscedasticity.

We run separate regressions for each trade measure and a joint one with effective tariff, import penetration and export orientation, we also include other variables related to the structure of the industry in order to test the robustness of the results. These variables are value-added[54], and the proportion of workers who earn less than the minimum wage. Additionally, we run one specification adding the industry-specific nominal exchange rate[55] to the joint specification. The set of regressions also contains year dummies aimed at capturing aggregate shocks that may have had some impact on all manufacturing industries.

Table 4.5 summarises the results of the fixed-effect with time dummies of the estimations for the wage premium as measured by the registered worker's dummy coefficient. Both effective and nominal tariffs have a (small) positive, but statistically insignificant coefficient. In contrast, the specification with import penetration as the only regressor shows a negative and significant coefficient (-0.0042). These results are robust to the inclusion of the other trade-related variables in column [5] and to industry characteristics (column[7]), as well as to the inclusion of the log of the industry-specific nominal exchange rate in column [6]. The inclusion of the latter variable aims at preventing the trade-related variables, particularly the import penetration ratio, from spuriously capturing the effect of different trends in the industry-specific exchange rates. The inclusion of industry-specific exchange rate does not affect either the sign or the statistical significance of the coefficients of the other trade-related variables. However, it slightly diminishes the negative effect of the import penetration ratio. The estimates range from -0.0033 to -0.0064[56]. The export orientation variable has a negative and statistically insignificant coefficient in all specifications. All additional controls have a positive effect on the wage premium, as expected, but they are never statistically significant.

These results suggest that trade liberalisation as measured by the import penetration ratio had a diminishing effect on the wage differential between registered and non-registered in the manufacturing sector. Overall, something between 10% and 14% of the 42% decrease in the wage gap in the manufacturing sector can be attributed to the 10% increase in the import penetration ratio. As for the other trade measures directly affected by the reform, the coefficients for nominal and effective tariff is correctly signed as we would expect based on our discussion in the

[54]The value-added was calculated based as the difference between total sales and total costs of the industries based on the data of the Industry Annual Survey (PIA) from IBGE, www.ibge.gov.br .

[55]The industry-specific exchange rate was calculated as the weighted average of the nominal exchange rates of countries accounting for more than 2% of industry exports in each year as in Revenga (1992). Due to the lack of good data on producer price indexes and the high inflation witnessed by the country until 1994, we decided not to calculate real exchange rates and to enter the industry-specific nominal exchange rate only in the specifications with time dummies.

[56]The estimates seem to be small because the explanatory variables entered the regressions as percentages.

dep. var=wage gap	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Effective Tariff	0.0008 [0.0007]				0.0007 [0.0007]	0.0008 [0.0007]	-0.0003 [0.0009]
Nominal Tariff		0.0009 [0.0013]					
Import Penetration			-0.0042 [0.0013]**		-0.0039 [0.0013]**	-0.0033 [0.0015]*	-0.0064 [0.0015]**
Export Orientation				-0.0011 [0.0016]	-0.0009 [0.0016]	-0.0004 [0.0017]	-0.0004 [0.0021]
Exchange Rate						-0.0096 [0.0089]	
% Workers earning less than mw							0.6548 [0.3437]
Value Added							0.0131 [0.0503]
Constant	0.4076 [0.0604]**	0.4152 [0.0844]**	0.4669 [0.0327]**	0.4711 [0.0323]**	0.4132 [0.0600]**	0.2252 [0.1876]	0.064 [1.1551]
N	170	17 0	170	170	170	170	153
Adj. R2	0.71	0.71	0.72	0.71	0.72	0.72	0.73
F test: industry	19.18	18.48	21.24	16.2	19.62	19.86	15.33
Prob >F	0	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Industry and time dummy variables not shown.
Note: As we do not have data on value-added for 1998, column [7] specification has only 153 observations.

Table 4.5: Wage Premium: Fixed Effect with Time Dummies (Contemporaneous Regressors)

last section, i.e., they would have a positive impact on the registered workers wage premium, but they are very small and not statistically significant.

The regressions for the proportion of registered workers in the pool of employees were weighted by the share of workers in the industry for each industry/year pair. The standard errors are Huber-White corrected. Table 4.6 shows the results for the fixed effect specifications with time dummies. Both effective and nominal tariffs have a negative effect on the proportion of registered workers for all specifications, and their coefficients are significant for most of them[57]. The coefficients for the import penetration ratio and for the export orientation are not significant, but while the import penetration ratio shows a negative impact in most specifications, the export orientation has a positive impact in all specifications[58]. The proportion of workers earning below the minimum wage is, as expected, negatively correlated with the proportion of registered workers. Including the industry-specific exchange rate does not change the coefficient (column [6]) of the the other variables.

The results in Tables 4.5 and 4.6 are somewhat puzzling. Whereas the import penetration ratio seems to have had a negative impact on the wage differential, it has not affected (significantly) the proportion of registered workers, despite its

[57]The coefficient on effective tariff is not significant at 5% level for the specification with additional controls, column [7].

[58]In fact, the export orientation ratio is significant in the joint specification.

dep. var=prop. reg	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Effective Tariff	-0.0003 [0.0001]**				-0.0004 [0.0001]**	-0.0004 [0.0001]**	-0.0002 [0.0001]
Nominal Tariff		-0.0003 [0.0002]					
Import Penetration			-0.0002 [0.0003]		-0.0003 [0.0002]	-0.0003 [0.0003]	0.0001 [0.0003]
Export Orientation				0.0006 [0.0004]	0.0008 [0.0004]*	0.0008 [0.0004]	0.0006 [0.0004]
Exchange Rate						-0.0001 [0.0021]	
% Workers earning less than mw							-0.2733 [0.0709]**
Value Added							-0.0021 [0.0105]
Constant	0.6976 [0.0116]**	0.6916 [0.0166]**	0.6713 [0.0087]**	0.6705 [0.0090]**	0.698 [0.0118]**	0.6962 [0.0421]**	0.8009 [0.2346]**
N	170	170	170	170	170	170	153
Adj. R2	0.97	0.97	0.97	0.97	0.97	0.97	0.97
F test: industry	289.89	280.66	227.76	302.16	203.25	203.19	46.03
Prob >F	0	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Industry and time dummy variables not shown.
Note: As we do not have data on value-added for 1998, column [7] specification has only 153 observations.

Table 4.6: Proportion of Registered Workers: Fixed Effect with Time Dummies (Contemporaneous Regressors)

negative sign in most specifications. The effective tariff has had no effect on the wage differential (despite its positive sign), but it has had a negative impact on the proportion of registered workers. Thus, industries most affected by the reduction in effective tariff were the same that witnessed an increase in the proportion of registered workers. Therefore, at least for the manufacturing sector, trade liberalisation may have had an impact on cutting the wage premia of registered workers, but if anything, it had the effect of increasing the proportion of registered workers in the most affected industries.

In order to check the robustness of these results, allowing for some delay in the adjustments to the new tariffs and to the more competitive environment, and also to avoid problems of simultaneity between import penetration ratio and export-orientation and the wage differential, we re-run equations (4.4) and (4.5) using lagged regressors rather than contemporaneous.

As for the results for the wage premium, Table 4.7 reveals that the main difference is that the coefficient of effective tariff is negative in most specifications, but again it is never statistically significant. The coefficients of export orientation turn out to be positive for most the fixed-effect specifications with time dummies, but they are not significant either. The coefficients for import penetration are negative and show a point estimate somewhat higher than the one with contemporaneous effect, ranging

dep. var= wage gap	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Effective Tariff	0.0002 [0.0008]				-0.0001 [0.0008]	-0.0000 [0.0008]	-0.0003 [0.0009]
Nominal Tariff		-0.0005 [0.0014]					
Import Penetration			-0.0048 [0.0016]**		-0.0061 [0.0016]**	-0.0055 [0.0019]**	-0.0064 [0.0022]**
Export Orientation				-0.0001 [0.0019]	0.0013 [0.0022]	0.0018 [0.0023]	0.0018 [0.0026]
Exchange Rate						-0.0078 [0.0079]	
% Workers earning less than mw							-0.1119 [0.6606]
Value Added							-0.0418 [0.0458]
Constant	0.3861 [0.0349]**	0.5788 [0.0979]**	0.4703 [0.0328]**	0.4693 [0.0321]**	0.3928 [0.0367]**	0.3931 [0.1704]*	1.2582 [1.0235]
N	153	153	170	170	153	153	119
Adj. R2	0.71	0.71	0.72	0.71	0.72	0.72	0.7
F test: industry	16.32	15.26	20.32	15	17.75	18.05	9.05
Prob >F	0	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Industry and time dummy variables not shown.
Note: As we do not have data on value-added for 1998, column [7] specification has only 153 observations.

Table 4.7: Wage Premium: Fixed Effect with Time Dummies (Lagged Regressors)

from -0.0048 to -0.0064. The inclusion of the lagged industry-specific exchange rate does not change this result (see column [6]).

As for the results for the proportion of registered workers, Table 4.8 shows that the use of a lagged specification instead of contemporaneous does not change the former results. Moreover, using the probability of registered job as the dependent variables as suggested in Goldberg and Pavcnik (2003) none of the trade variables in any of the specifications are statistically significant.[59]

Altogether, the results for both set of equations suggest that trade reforms as measured by the import penetration ratio had a diminishing effect in the wage differential between registered and non-registered workers. However, they also suggest that trade liberalisation did not play any role in the increase of the proportion of non-registered workers in the manufacturing sector. If anything, the effective tariff seems to have a negative impact on the proportion of registered workers and no role in the diminishing wage differential between the two groups. Nevertheless, the narrowing effect of the increase in the import penetration ratio on the wage gap seems to be more robust than the effect of effective tariff on the proportion of registered workers. Using lagged regressors does not change the results for the wage differential, and only slightly changes the results for the proportion of registered workers when more controls are added. Moreover, it is important to keep in mind that the

[59]The results are available upon request.

dep. var=prop. reg	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Effective Tariff	-0.0003 [0.0001]*				-0.0003 [0.0001]*	-0.0003 [0.0001]*	-0.0001 [0.0002]
Nominal Tariff		-0.0004 [0.0003]					
Import Penetration			-0.0001 [0.0003]		-0.0001 [0.0003]	0.0000 [0.0005]	0.0003 [0.0004]
Export Orientation				0.0004 [0.0005]	0.0001 [0.0004]	0.0002 [0.0005]	0.0006 [0.0005]
Exchange Rate						-0.0018 [0.0022]	
% Workers earning less than mw							-0.349 [0.0653]**
Value Added							0.0183 [0.0114]
Constant	0.6419 [0.0087]**	0.6435 [0.0089]**	0.6716 [0.0088]**	0.6714 [0.0088]**	0.6419 [0.0088]**	0.656 [0.0451]**	0.3663 [0.2617]
N	153	153	170	170	153	153	119
Adj. R2	0.97	0.97	0.97	0.97	0.97	0.97	0.97
F test: industry	266.07	244.75	226.28	310.58	193.33	194.34	42.1
Prob >F	0	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Industry and time dummy variables not shown.
Note: As we do not have data on value-added for 1998, column [7] specification has only 153 observations.

Table 4.8: Proportion of Registered Workers: Fixed Effect with Time Dummies (Lagged Regressors)

effective tariff does not take into account the lifting of important trade barriers that may have affected the actual degree of competitiveness in the industry. This effect is captured, at least partially, by the import penetration ratio.

4.5.2 Spillover Effects of Trade Liberalisation

In order to check whether the process of trade liberalisation had any impact beyond the supposedly direct effect on the manufacturing labour market, we run some fixed-effect models exploiting the different composition of the manufacturing sector in different regions of the country. The idea is to assess whether or not the economy-wide narrowing of the wage gap between formal and informal workers can be attributed to the trade liberalisation process[60]. To implement that, we use the wage gap in different regions as the dependent variables and the trade measures as the regressors. As there is no regional variation in the trade measures since these data are only available at national level, we use the following strategy: we take the share of workers in each of the 17 industries in 1987 within the following eight regions: metropolitan Southeast, non-metropolitan Southeast, metropolitan

[60]Notice that the impact of the trade liberalisation on wages as discussed in section 4.3.1 applies to the whole economy and not only to the manufacturing sector.

South, non-metropolitan South, metropolitan Northeast, non-metropolitan Northeast, North and Midwest, and multiply this share by the effective and nominal tariffs, by the import penetration ratio and by export orientation ratio for each industry, and calculated the regional weighted average for each of these variables. These modified trade-related variables would indicate the intensity with which each regional labour market must have been affected by trade liberalisation. The source of variation now comes from the different degrees of protection between industries and from their regional dispersion within the country.

One caveat of this procedure is that trade liberalisation is likely to have led to a reallocation of the industries within the country. In order to face a more competitive environment firms are likely to move to places where their costs can be lowered. This is one of the reasons why we choose to use a fixed weight - the industry structure in 1987.

It is important to bear in mind that the results for this set of equations are not comparable to the previous one, since now we are looking at the effect of trade liberalisation on the economy-wide wage differential between registered and non-registered workers and on the economy-wide proportion of registered workers in the economy, and not only in the manufacturing sector. Equations (4.3), (4.4), and (4.5) may be rewritten as:

$$w_{irt} = \alpha_3 + \Lambda_{rt} Z_{irt} + \beta_{rt} Reg_{irt} + \varepsilon_{irt} \tag{4.6}$$

where w_{irt} is the log of the real hourly-wage for the individual i in region r and year t and Λ is a vector of coefficients correspondent to the following independent variables Z: gender, education (6 groups)[61], experience, experience squared, and β_{rt} is the coefficient for the dummy variable *Reg* that indicates whether the individual is a registered worker or not.

$$\beta_{rt} = \alpha_4 + \pi TM_{rt} + \mu_r + \theta_t + \varepsilon_{rt} \tag{4.7}$$

where *TM* stands for trade measures variables as defined above, μ_r are region dummies and θ_t, time period dummies and ε_{rt} is assumed to be a white noise.

$$Preg_{rt} = \alpha_5 + \kappa TM_{jt} + \mu_r + \theta_t + \varepsilon_{rt} \tag{4.8}$$

where *Preg* is the proportion of registered workers in region r and year t and the remaining variables are as stated in equation (4.5).

[61] The education groups are the same ones defined in chapter 2.

dep. var=wage gap	[1]	[2]	[3]	[4]	[5]	[6]
Regional Effective Tariff	-0.005616				-0.006354	-0.003711
	[0.003032]				[0.003841]	[0.003809]
Regional Nominal Tariff		-0.00547				
		[0.005242]				
Regional Import Penetration			0.005169		0.008033	-0.00047
			[0.010112]		[0.008984]	[0.008203]
Regional Export Orientation				-0.005762	0.001539	-0.000944
				[0.012008]	[0.015678]	[0.017714]
Proportion of Workers in Manufacturing						-0.217793
						[0.345291]
% Workers earning less than mw						0.639823
						[0.291847]*
Constant	0.659821	0.650932	0.622713	0.626998	0.662332	0.648209
	[0.033614]**	[0.039330]**	[0.028562]**	[0.029977]**	[0.034016]**	[0.373064]
N	73	73	73	73	73	73
Adj. R2	0.91	0.91	0.9	0.9	0.91	0.91
F test: region	63.34	63.41	61.34	60.31	58.98	14.06
Prob >F	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Region and time dummy variables not shown.

Table 4.9: Regional Wage Premium: Fixed Effect with Time Dummies (Contemporaneous Regressors)

The fixed-effect with time dummies versions of equation (4.7) are weighted by the inverse of the sampling variance of the dependent variable, whereas the fixed-effects with time dummies versions of equation (4.8) are weighted by the share of workers in each region. We also include additional controls in one of the specifications in order to check the robustness of the results. The controls are: proportion of workers employed in the manufacturing sector in each region, and the proportion of worker who earn less than the minimum wage.

Table 4.9 shows the results for the estimation of equation (4.7) with year and industry dummies using regional variation to identify the impact of trade related variables on the entire labour market. The effective and nominal tariffs have a negative and insignificant effect on wage differential between registered and non-registered workers. The sign of the import penetration ratio and of the export orientation are very sensitive according to the specification, but they are never significant. The only variable that is statistically significant in the specifications showed in Table 4.9 is the proportion of workers earning less than the minimum wage that has a positive impact on the wage premium of registered workers. Therefore, it seems that the trade-related variables did not have any impact on the behaviour of the wage differential in terms of its regional variation.

As for the proportion of registered workers in the pool of employees, Table 4.10

dep. var=prop. reg	[1]	[2]	[3]	[4]	[5]	[6]
Regional Effective Tariff	0.003072				0.001908	0.000015
	[0.002121]				[0.001979]	[0.001120]
Regional Nominal Tariff		0.003283				
		[0.002945]				
Regional Import Penetration			-0.013051		-0.017194	-0.001762
			[0.005686]*		[0.005023]**	[0.002364]
Regional Export Orientation				0.01275	0.013671	0.006479
				[0.004933]*	[0.004381]**	[0.003039]*
Proportion of Workers in Manufacturing						0.839754
						[0.125824]**
% Workers earning less than mw						-0.375648
						[0.073687]**
Constant	0.633378	0.637003	0.656274	0.645894	0.635496	-0.023001
	[0.012382]**	[0.019419]**	[0.012439]**	[0.012284]**	[0.011026]**	[0.126030]
N	73	73	73	73	73	73
Adj. R2	0.97	0.97	0.97	0.97	0.98	0.99
F test: region	395.2	409.92	426.82	527.53	576.04	141.3
Prob >F	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Region and time dummy variables not shown.

Table 4.10: Regional Proportion of Registered Workers: Fixed Effect with Time Dummies (Contemporaneous Regressors)

shows that the import penetration ratio coefficients are negative for three specifications, but it is not significant in the specification with additional controls. The export orientation coefficients, on the other hand, have a positive and significant effect on the proportion of registered workers. Both nominal and effective tariff have a positive, but insignificant effect. As for the other controls, it is quite remarkable the positive impact of the proportion of workers employed on the manufacturing sector and the negative effect of the proportion of workers earning below the minimum wage on the proportion of registered workers. If anything, we could only say from these results that there is some weak evidence that import penetration ratio is negatively correlated with the proportion of registered workers and some strong evidence that the export orientation ratio is positively correlated with the proportion of registered workers.

The results above suggest that import penetration ratio affects negatively the proportion of registered workers in the entire economy when we use regional variation to identify the impact of trade liberalisation. Regions with industries more affected by the higher import penetration were the regions that witnessed a higher decrease in the proportion of registered workers. Similarly, regions with industries with a high export orientation ratio witness an increase in the proportion of registered workers.

dep. var=wage gap	[1]	[2]	[3]	[4]	[5]	[6]
Regional Effective Tariff	-0.006393				-0.006038	0.00144
	[0.005678]				[0.006716]	[0.007599]
Regional Nominal Tariff		-0.004922				
		[0.008077]				
Regional Import Penetration			-0.034032		-0.02887	-0.01803
			[0.025302]		[0.026758]	[0.027799]
Regional Export Orientation				-0.007305	0.004199	-0.010641
				[0.011589]	[0.017899]	[0.022503]
Proportion of Workers in Manufacturing						-0.163564
						[0.371722]
% Workers earning less than mw						0.820167
						[0.426472]
Constant	0.69069	0.687941	0.645335	0.639479	0.694627	0.489913
	[0.042675]**	[0.042119]**	[0.027623]**	[0.027924]**	[0.044029]**	[0.458521]
N	59	59	65	65	59	59
Adj. R2	0.9	0.9	0.9	0.9	0.9	0.91
F test: region	47.77	48.91	45.46	51.56	32.97	10.13
Prob >F	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Region and time dummy variables not shown.

Table 4.11: Regional Wage Premium: Fixed Effect with Time Dummies (Lagged Regressors)

However, unlike the specifications for the manufacturing sector, there has been no evidence that trade-related measures have affected the wage differential between registered and non-registered workers under this identification strategy.

The re-estimation of (4.7) and (4.8) using lagged regressors instead of contemporaneous yields very similar results for the impact of the trade-related measures on the wage differential. Table 4.11 shows that there is no indication of any impact of trade liberalisation on the wage differential. However, for the impact of trade-related measures on the proportion of registered workers the story is somewhat different. Table 4.12 shows that the import penetration ratio has now a positive, but not significant effect on the proportion of registered workers. Furthermore, the positive effect of the export orientation ratio becomes insignificant.

All in all we find evidence that the increase in the import penetration ratio seems to have lowered the wage differential within the manufacturing sector. As for the proportion of registered workers in the manufacturing sector, we find some evidence that contemporaneous effective tariffs are negatively correlated with the proportion of registered workers. We do not find any evidence of spillover effect of trade liberalisation on the economy-wide wage gap. But there is some weak evidence that the proportion of registered workers in the entire (regional) economy

dep. var: prop. reg	[1]	[2]	[3]	[4]	[5]	[6]
Regional Effective Tariff	0.003685				0.002525	0.000779
	[0.002967]				[0.003738]	[0.001426]
Regional Nominal Tariff		0.002596				
		[0.003509]				
Regional Import Penetration			0.004658		0.005012	0.006171
			[0.011899]		[0.012131]	[0.005171]
Regional Export Orientation				0.002155	0.003661	-0.004568
				[0.005033]	[0.007387]	[0.005023]
Proportion of Workers in Manufacturing						0.900756
						[0.123634]**
% Workers earning less than mw						-0.337007
						[0.079362]**
Constant	0.621272	0.623155	0.641534	0.641249	0.620169	-0.099816
	[0.009507]**	[0.009352]**	[0.009311]**	[0.009682]**	[0.010300]**	[0.147417]
N	59	59	65	65	59	59
Adj. R2	0.98	0.98	0.98	0.98	0.98	1
F test: region	331.86	333.74	363.71	357.25	234.34	237.13
Prob >F	0	0	0	0	0	0

Robust standard errors in brackets. * Significant at 5%; ** significant at 1%.
Region and time dummy variables not shown.

Table 4.12: Regional Proportion of Registered Workers: Fixed Effect with Time Dummies (Lagged Regressors)

was negatively affected by the increase in import penetration ratio. But again this result does not hold when we enter lagged regressors.

4.5.3 Pseudo-cohort Approach

This last approach aims at putting together the variation exhibited by both tradable and non-tradable sectors in order to assess the impact of trade liberalisation. Therefore, besides the information (variation) due to the heterogeneous impact on the manufacturing sector (tradable sector) of the trade-related variables, we would be also using the information (variation) on the non-traded sector as a comparison group, which was not directly affected by the trade reform to identify the overall effect of trade liberalisation on the wage premium for registered workers and for its proportion in the pool of employees.

In the absence of a panel data set that covers most of the country, we decided to build a pseudo-panel from the repeated cross sections that we have been using so far. In order to get a better view of what is going on, the unit of analysis is no longer the industry, but industry-cohort cells. The use of cohort allow us to control among other things for the quality of education and for the possibility of a life-cycle component in the distribution of workers into formal and informal sector. There

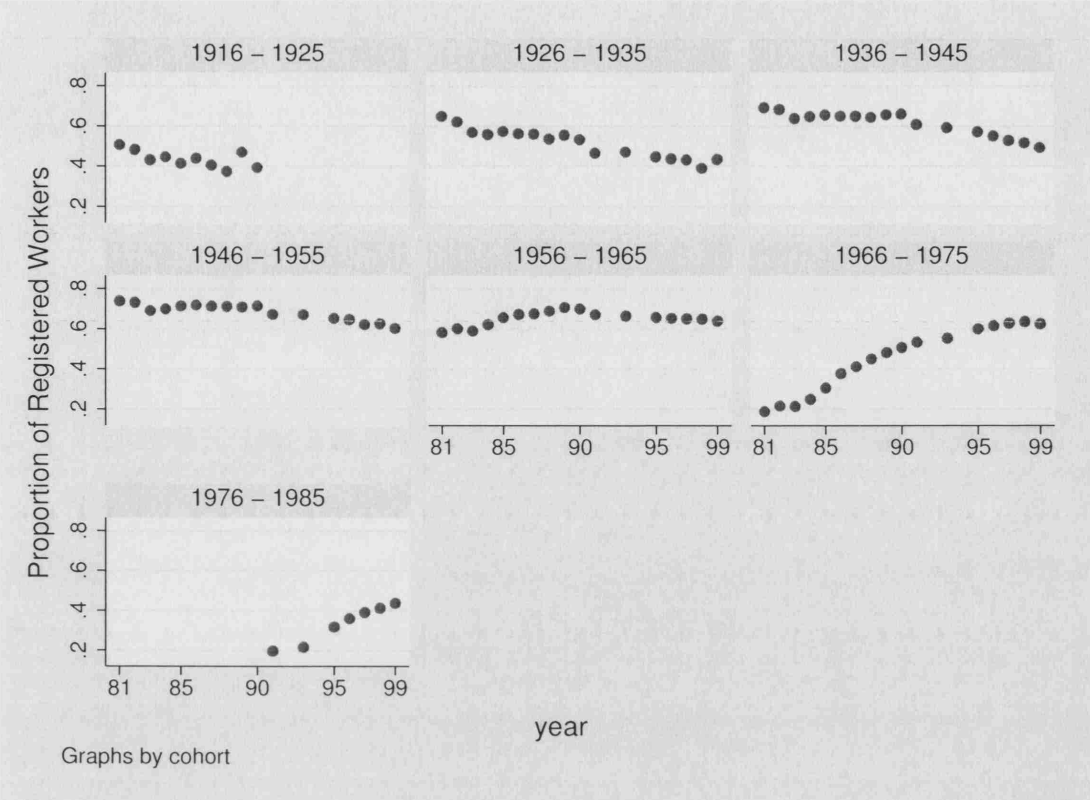

Figure 4.8: Proportion of Registered workers by cohort (1981 - 1999)

is some evidence that younger workers can start at the informal sector and later on as they accumulate human capital (experience) they move to the formal sector. We divided the sample in 7 cohorts of 10-year length and 30 industry-cells. From those 30 industry-cells, 19 are classified as tradable sectors[62] and 11 as non-tradable sectors. The latter comprises mainly activities such as construction sector, lodging, restaurants, personal services, and productive services.

Figure 4.8 show the proportion of registered workers for each cohort cell between 1991 and 1999. There is lots of variation between cohort-cells probably due to life cycle changes on the employment composition. The panel is unbalanced due to the fact that persons from the first cohort (born between 1916 and 1925) do not show up after 1990 (we drop people older than 65 from the sample) and persons from the last cohort (born between 1976 and 1985) only show up in the panel from 1990 onwards. Moreover, when some cohorts are still too small (either full of young people or older people), they have missing observations for some industries. We use data from 1981

[62]Besides the 17 industries used previously, we also add agriculture and the mineral extracting industry to the group of tradable sector. As there is no data on import penetration for the agriculture sector, we also run some specifications without it.

to 1998[63], thus for complete industry-cohort cell observations we should observe 16 observations for each cell. Table 4.13 shows the number of years observed for each industry-cohort cell. The maximum number of observations in our estimations is 2867[64]. However, more worrying than the missing cells is whether or not the number of observations that have generated the mean value cells for each industry-cohort and for each year is large enough, so that we can ignore the correction put forward by Deaton (1985). As Deaton points out, if one intends to get an unbiased estimate, then one must take into account the variability of the calculated cell mean in the estimation of the parameters. Nevertheless, several studies have shown that when the number of observations that have generated the mean cell is higher than 100, the correction advocated by Deaton does not make much difference. From the 2867 mean cell observations, 56% (1602) were generated from 100 or more individual observations. For now, we decide to run the regressions without using the correction, but we weight each cell mean observation by the number of observations that have generated it.

In order to investigate the impact of trade liberalization on wage differential and on the proportion of registered workers we adopt several different strategies, always using fixed-effect model with time dummies, where the unit of analysis is the industry-cohort cell.

First, we will approximate the impact of trade liberalisation by a binary dummy (*lib92*) that assumes value 1 from 1992 onwards and 0 otherwise. We use this definition because the bulk of the trade liberalisation process occurred between 1990 and 1992. This strategy will allow us to use the whole sample period from 1981 to 1998. The sample period in this case is different from the one in the previous section because the trade-related variables that we have used there are only available from 1987 onwards. The use of interactions between the binary *lib92* and another binary variable named *trade* (trade=1 if industry is tradable, and 0 otherwise) will allow us to infer whether the wage of workers in the tradable sector had a different behaviour when compared to the wages of workers in the non-tradable sector after the trade liberalisation. Similar interactions will be used to assess the impact on the wage premium for registered workers. In this case, we interact *lib92* with the proportion of registered workers in the specific industry-cohort cell (*preg*).

Second, we will add to the regression the trade-related variables: effective tariff and import penetration ratio as we did in the previous sections. In this case, the

[63] As mentioned in Chapter 2 there is no data for 1991 and 1994.

[64] This corresponds to the specifications that use data for all years and all sectors (including agriculture).

	Cohort							Total
Industry	1	2	3	4	5	6	7	
Tradable								
1. Agriculture	10	16	16	16	16	16	7	97
2. Ext.mineral	10	16	16	16	16	16	7	97
3. Non-metalic	10	16	16	16	16	16	7	97
4. Metallic	9	16	16	16	16	16	7	96
5. Mechanic	10	15	16	16	16	16	7	96
6. Electric and electronics	10	15	16	16	16	16	7	96
7. Vehicles	10	16	16	16	16	16	7	97
8. Wood	10	16	16	16	16	16	7	97
9. Paper	10	16	16	16	16	16	7	97
10. Rubber	8	14	16	16	16	16	7	93
11. Shoes	10	16	16	16	16	16	7	97
12. Chemicals	9	16	16	16	16	16	7	96
13. Oil refining	6	13	16	16	16	16	6	89
14. Pharmaceutical	9	15	16	16	16	16	7	95
15. Plastics	8	16	16	16	16	16	7	95
16. Textiles	9	16	16	16	16	16	7	96
17. Clothes	10	16	16	16	16	16	7	97
18. Food	10	16	16	16	16	16	7	97
19. Other manufacturing	9	15	16	16	16	16	7	95
Non-Tradable								
20. Construction	10	16	16	16	16	16	7	97
21. Industrial services	10	16	16	16	16	16	7	97
22. Commerce	10	16	16	16	16	16	7	97
23. Financial institution	9	16	16	16	16	16	7	96
24. Transport	10	16	16	16	16	16	7	97
25. Communications	9	16	16	16	16	16	7	96
26. Family services	10	16	16	16	16	16	7	97
27. Company services	10	16	16	16	16	16	7	97
28. Other non-traded services	10	16	16	16	16	16	7	97
29. Rentals	10	16	16	16	16	16	7	97
30. Petrol and gas	6	11	16	16	16	12	2	79
Total	281	466	480	480	480	476	204	2867

Table 4.13: Number of Years Observed for Each Industry-cohort Cell

sample period will be reduced to the period 1987-1998; however we will be using additional variation within the manufacturing sector in order to estimate the effect of the changes in trade-related measures. In order to standardize the trade-related measures so that we can put together the data in meaningful way for both tradable and non-tradable sectors, we follow Dickerson et al. (2001) and Arbache et al. (2004) and use the following transformations[65]:

Openness=exp(-effective tariff) if the industry is in the tradable sector and *Openness*=0 otherwise.;

Closedness=exp(-import penetration ratio) if the industry is in the tradable sector and *closedness*=1 otherwise.

These transformations enable us to restrict the *openness* and *closedness* measures inside the range [0,1]. For the *openness* measure, the higher the effective tariff the closer the index will be to 0; whereas for the *closedness* measure, the lower the import penetration ratio, the closer the index will be to 1.

Besides the trade-related variables, i.e., *lib92* dummy, *openness* and *closedness* indices, we also control for human capital variables - 6 groups of educations (illiterate as the base category), experience and experience squared – gender and registered status. The registered status variable will give us a measure of the wage differential between registered and non-registered workers[66].

The Impact on Wages

Table 4.14 shows the results of the fixed-effect specifications with time dummies. The coefficients on column [1] suggest a wage premium of 43%[67] for registered workers. As for the impact of trade liberalisation, the coefficient of *lib92* and of its interactions with *trade* show that after 1992 tradable sector workers experienced a fall of 13% in their wage relative to their counterparts in the non-tradable sector[68]. Column [2] shows that registered workers in the tradable sector earn less than registered workers in the non-tradable sector *reg*trade.* Column [3] shows that after 1992 the wage differential in favour of registered workers fell by 17%. Columns [4] to [6] bring

[65] Behrman et al. (2001) use the maximum of their *index of openness* (average nominal tariff*standard deviation of nominal tariff) as the normalisation factor for their sample of countries. They argue that this would be a more credible measure than a [0,1] scale. However, we cannot use such scale since by definition our index measures assume extreme values (either 0 or 1) for the non-tradable industries.

[66] Notice that this variable for each industry-cohort cell is not a binary dummy (1/0), but the proportion of registered workers on that cell.

[67] Calculated as [100*(exp(0.38)–1)].

[68] Calculated as 100*[exp(-0.1537+0.0128) -1].

the results for these same specifications, but excluding the agricultural sector[69]. The previous results do not change much. The estimated wage premium for registered workers is higher, 95%. After 1992, the wage for workers in the tradable sector fell by 11%. Registered workers in the tradable sector earned less than registered workers in the non-tradable sector. But the most striking result is the fall in the wage of registered workers by 55% after 1992. It is clear from these results that after the trade liberalisation, the wage of workers in the tradable sector fell relatively to the wage of workers in the non-tradable sector and that the wage differential in favour of registered workers fell sharply.

So far we discussed the different response of the tradable and non-tradable sector wages to changes before and after the trade liberalisation in order to identify its impact on wage differential between registered and non-registered workers. To add more variability in the way that the trade liberalisation may have affected the wage of workers in different industries, we introduce the variables *openness* and *closedness* and their interactions with the proportion of registered workers in order to capture the different intensity that the reform affected different tradable industries. Hence, besides variation between tradable and non-tradable sectors, we add the variation within the tradable sector to identify the impact of trade liberalisation on the wage differential between registered and non-registered workers. Columns [7] to [10] in Table 4.14 show the results for the restricted sample that covers the period 1987-1998 and excludes the agricultural sector. This smaller sample period changes a bit the magnitude, but not the sign of the previous result: there is a positive wage premium for registered workers, but this wage premium is lower in the tradable sector. The *openness index* added in column [8] is negative and statistically significant. When we add the interaction between the openness measure and the proportion of registered workers (*openness*reg*), both variables (*openness* and *openness*reg*) turn out to be insignificant, but with opposite signs. The coefficient on the interaction (*openness*reg*) is relatively high and negative, whereas the coefficient on *openness* is small and positive. The fall in effective tariffs seems to have affected negatively the wages and this effect is mainly borne by registered workers. The *closedness* index display a positive and significant coefficient when entered without interaction as in Column [9]. Thus, increases in the import penetration ratio leads to a fall in wages. Column [10] shows that this effect is stronger the higher the proportion of registered workers in the sector.

[69]Note that the sample size was reduced to 2.770 observations.

	1981 -1998						1987 - 1998			
	Including agricultural sector			Excluding agricultural sector			Excluding agricultural sector			
dep. var: lwh	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Reg	0.382	0.578	0.566	0.671	0.715	0.662	0.928	0.932	0.879	0.881
	[0.035]**	[0.039]**	[0.041]**	[0.039]**	[0.041]**	[0.037]**	[0.058]**	[0.058]**	[0.058]**	[0.058]**
lib92	0.013	0.059	0.164	-0.038	-0.029	0.263				
	[0.057]	[0.056]	[0.059]**	[0.061]	[0.061]	[0.058]**				
(lib92) * trade	-0.154	-0.126		-0.105	-0.101					
	[0.007]**	[0.008]**		[0.008]**	[0.008]**					
(lib92) * reg			-0.198			-0.444				
			[0.018]**			[0.018]**				
Reg * trade		-0.513	-0.868		-0.179	-0.120	-0.483	-0.470	-0.457	-0.544
		[0.051]**	[0.049]**		[0.055]**	[0.050]*	[0.082]**	[0.083]**	[0.081]**	[0.088]**
Openness							-0.085	0.069		
							[0.036]*	[0.118]		
Openness * reg								-0.284		
								[0.209]		
Closedness									0.168	-0.212
									[0.031]**	[0.156]
Closedness * reg										0.5248
										[0.2109]*
Constant	-1.426	-1.258	-1.093	-1.291	-1.248	-1.297	-1.424	-1.433	-1.406	-1.437
	[0.096]**	[0.095]**	[0.097]**	[0.095]**	[0.096]**	[0.089]**	[0.160]**	[0.160]**	[0.159]**	[0.159]**
N	2867	2867	2867	2770	2770	2770	1731	1731	1731	1731
R2	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
F test: id dummy	25.62	26.12	25.1	30.67	30.13	30.35	21.68	21.7	22.06	22.04

Standard errors in brackets. * Significant at 5%; ** significant at 1%
Human capital variables, gender, and industry-cohort and time dummies not shown.

Table 4.14: Pseudo-Cohort: Fixed Effects with Time Dummies for log Hourly Wage

All in all, it seems that all trade-related measures in the fixed-effect specifications (*lib92* dummy, *openness* index and *closedness* index) point to lower wages in the sectors most affected by the trade liberalisation, and among the most affected it seems that the registered workers were the ones who suffered most. Altogether, these results suggest that, in fact, trade liberalisation had a role in narrowing the wage gap between registered and non-registered workers. It is noteworthy that the impact of the import penetration ratio in this framework is similar to the one we found within the manufacturing sector in our first strategy.

The impact on the Proportion of Registered Workers

The results for the fixed-effect model with time dummies for the proportion of registered workers in Table 4.15 show that the proportion of men and of more educated people is positively correlated with the proportion of registered workers regardless of the specification in use. Columns [2] to [4] reveal that both tradable and non-tradable sectors had an increase in the proportion of registered workers after 1992, and the tradable sector grew by a lower amount in the sample without the agricultural sector[70]. However, when the sample is restricted to the period 1987 to 1998 and without the agricultural sector, as in Column [5], we see no change in the non-tradable sector after 1992 and a decrease in the proportion of registered workers in the tradable sector. Columns [6] and [7] show the results including *openness* and *closedness* indices. The coefficient of the openness index is positive and statistically significant (0.0414), implying that the higher the *openness* index the higher the proportion of registered workers. As for the coefficient of the *closedness* index, it is also positive and statistically significant[71].

Again, there is no strong evidence that trade liberalisation has triggered an increase in the proportion of non-registered workers. The fixed effect models with time dummies revealed a positive effect of the fall in tariffs on the proportion of registered workers and at the same time a negative effect of the increase in the import penetration ratio. Thus, the fall in tariff would have led to an increase in the proportion of registered workers whereas the increase in the import penetration ratio would have led to a decrease in the proportion of registered workers.

[70]It is worth noting that unlike the "urban sectors", the agricultural sector has experienced an increase in the proportion of registered workers from 1981 to 1999. See Chapter 2 for more details on that.

[71]This latter result is in line with the one we found using the Spearman rank correlation.

	1981 1998				1987 1998		
	Including agricultural sector		Excluding agricultural sector		Excluding agricultural sector		
dep. var: prop. reg	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Some elementary	-0.058	-0.060	0.180	0.198	-0.223	-0.253	-0.247
	[0.0505]	[0.0504]	[0.0474]**	[0.0465]**	[0.0601]**	[0.0606]**	[0.0603]**
Complete elementary	0.477	0.475	0.515	0.495	0.108	0.100	0.095
	[0.0405]**	[0.0405]**	[0.0374]**	[0.0367]**	[0.0540]*	[0.0546]	[0.0543]
Complete primary	0.835	0.836	0.881	0.880	0.510	0.493	0.497
	[0.0527]**	[0.0527]**	[0.0482]**	[0.0472]**	[0.0627]**	[0.0633]**	[0.0629]**
Complete secondary	0.74	0.74	0.73	0.72	0.37	0.36	0.35
	[0.0552]**	[0.0552]**	[0.0510]**	[0.0500]**	[0.0690]**	[0.0698]**	[0.0694]**
Complete college	0.399	0.409	0.405	0.381	0.122	0.128	0.109
	[0.0683]**	[0.0683]**	[0.0629]**	[0.0617]**	[0.0846]	[0.0854]	[0.0851]
exp	0.002	0.003	-0.002	-0.003	0.009	0.009	0.009
	[0.0019]	[0.0019]	[0.0019]	[0.0019]	[0.0029]**	[0.0029]**	[0.0029]**
exp^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	[0.0000]**	[0.0000]**	[0.0000]**	[0.0000]**	[0.0000]**	[0.0000]**	[0.0000]**
sex	0.128	0.121	0.095	0.118	0.177	0.148	0.155
	[0.0263]**	[0.0264]**	[0.0231]**	[0.0228]**	[0.0301]**	[0.0301]**	[0.0299]**
lib92	0.165	0.039	0.268	0.130	0.000		
	[0.0315]**	[0.0088]**	[0.0315]**	[0.0156]**	[0.0000]		
lib92 * trade		0.010		-0.041	-0.026		
		[0.0040]**		[0.0040]**	[0.0043]**		
Openness						0.041	
						[0.0172]*	
Closedness							0.068
							[0.0148]**
Constant	0.126	0.129	0.177	0.175	0.340	0.365	0.363
	[0.0533]*	[0.0532]*	[0.0499]**	[0.0489]**	[0.0761]**	[0.0768]**	[0.0764]**
N	2867	2867	2770	2770	1731	1731	1731
R2	0.98	0.98	0.97	0.97	0.98	0.98	0.98
F test: id dummy	52.2	40.64	62.15	34.68	24.92	44.6	41.79

Standard errors in brackets. * Significant at 5%; ** significant at 1%
Industry-cohort and time dummies not shown.

Table 4.15: Pseudo-Cohort: Fixed-Effect with Time Dummies for the Proportion of Registered Workers

4.6 Conclusion

The fall in the wage differential between registered and non-registered workers as well as the fall in the proportion of registered workers are two stylised facts - as seen in Chapter 2 - of the Brazilian labour market in the 1990's. In this chapter, we investigated whether or not trade liberalisation had a role in this process using three strategies: 1) variation in the manufacturing sector; 2) (artificial) regional variation and 3) industry-cohort variation.

Overall, we found evidence that the trade liberalisation process is behind the fall in the wage gap in the manufacturing sector. It seems that rents that went to registered workers were cut due to the more competitive environment in the economy. This result was found in both the first strategy and the third strategy used in this chapter. The first strategy used only within-manufacturing industry variation and revealed a negative effect of import penetration on the wage gap. The third strategy used the non-tradable sector as a comparison group in addition to the variation within the manufacturing sector. In this latter case, both the fall in tariffs as measured by the *openness* index and the increase in import penetration ratio as measured by the *closedness* index led to a fall in the wage gap between formal and informal workers. The use of the non-tradable sector as a comparison group in the third strategy can be justified by the fact that we failed to find convincing evidence that trade liberalisation had any impact on the economy-wide fall in the wage gap according to the second strategy, that used regional variation in the industry dispersion to identify the effects of trade liberalisation in the entire economy.

The fall in the proportion of registered workers, however, does not seem to be correlated with trade liberalisation[72]. None of the strategies yielded robust evidence that the trade measures were correlated with the fall in the proportion of registered workers. The evidence found in the first strategy that effective tariff led to an increase in the proportion of registered workers in the manufacturing sector was not robust to the use of lagged regressors in the specification with additional controls. Similarly, the evidence yielded by the second strategy that the increase in the import penetration ratio led to a fall in the proportion of registered workers in the entire labour market is not robust to the use of lagged regressors. According to the third strategy, the fall in tariffs led to a higher proportion of registered workers[73], whereas the increase in the import penetration ratio led to a fall in the proportion of registered workers.

[72]This result is in line with the results of Goldberg and Pavnick (2003) for Brazil using a different data source, the Monthly Employment Survey (PME).

[73]Similar results was found in the first strategy.

Therefore, the weak evidence for the effect of trade liberalisation on the proportion of registered workers suggests that the fall in the proportion of registered workers was due to macroeconomic factors or institutional changes that had affected in a homogenous way regions and industries within the country.

A possible caveat of the approach adopted in this chapter is the fact that we do not look at the effects of trade liberalisation on the mobility pattern of the working population. Focusing on the effect on the proportion of registered workers and on the wage differential gives us only a partial account of the possible effects of those changes in the labour market. An interesting topic of research would be to analyse what happened to the size of the queue for formal jobs after the trade liberalisation. Unfortunately, we do not have surveys that contain a question similar to the one used in Chapter 3 to identify informal workers in the queue for formal jobs for periods after the trade liberalisation. However, as the wage differential is one of the main determinants of the queue for formal jobs, it is reasonable to infer that the size of the queue may have diminished after trade liberalisation due to the narrowing of the gap between formal and informal workers.

Chapter 5

Minimum wage and the Informal Sector in Brazil

5.1 Introduction

The minimum wage has been singled out in the literature as one of the causes of segmentation[1] between formal and informal sector workers in developing countries (Behrman,1999). The idea is that the informal sector is not covered by the minimum wage policy or that compliance is not enforced is that sector. Thus minimum wage hikes would lead to an increase in the informal sector because small firms that could not afford the wage increase would start hiring workers with a wage lower than the minimum wage[2] in the informal sector.

As seen in Chapter 2, the degree of non-compliance with the minimum wage decreased a lot from 1981 to 1999 in Brazil. In the informal sector the non-compliance fell from 61% to 35% in the informal sector and from 7% to 1% in the formal sector. Possibly this happened due to the fall in the real value of the minimum wage over this period. The high level of non-compliance in the informal sector has led some to believe, that minimum wage hikes were irrelevant for the determination of the wage of those workers. However, some models on the effect of the minimum wage in the covered and non-covered sectors challenged this common sense. Welch (1974) develops a model with two sectors: one covered by the minimum wage and the other not covered. In his model, minimum wage hikes lead to both an increase in the non-covered sector and a reduction in the participation rate, since workers for whom the new (lower) equilibrium wage in the non-covered sector is below their reservation

[1]See for instance Rauch's (1991) model discussed in chapter 2.

[2]We are assuming that the cost of non-compliance is higher for large firms, since they are more easily caught by inspection than small firms.

wage would withdraw from the labour market. Mincer (1976) extends this model to allow the existence of unemployment. In his model the effect of minimum wage hikes on the non-covered sector will depend whether the turnover rate is higher or lower than the elasticity of demand in the protected sector. If it is higher, than the non-covered wage will increase, as workers would queue for covered jobs. If its lower, the wage in the non-covered sector would fall, as workers migrate to this sector.

Thus, workers affected by minimum wage hikes can move to different states after the increase. Minimum wage hikes can affect transitions to the informal sector, to unemployment, to inactivity and even transitions from inactivity to unemployment. What happens with a worker's mobility pattern after a minimum wage hike is mainly an empirical question.

Besides having a sizable informal sector, the Brazilian labour market has special features that make the study of the impact of the minimum wage on the labour market even more complex. First, the formal sector during the 1980's and the first half of the 1990's followed strict rules of wage indexation determined centrally by the federal government. In addition, during this period there was a high concentration of formal sector workers around multiples of the minimum wage, so that the spike on the value of minimum wage in the distribution of wages was not unique. Thus, minimum wage hikes were quite likely to have spillover effects in the wage distribution of formal sector workers. For the informal sector, such spikes were not observed in the 1980's. However, since 1995 there has been an increasing indexation of their wages to the minimum wage. Furthermore, it has been observed that a higher proportion of informal sector workers have received wages increases equal to the increase of the minimum wage than formal sector workers. Therefore, the minimum wage policy in Brazil is likely to have a direct effect on informal sector workers, at least, in the second half of the 1990's, in addition to the indirect effect pointed out in models like the ones developed by Welch (1974) and Mincer (1976).

In this chapter we will investigate the effect of the minimum wage hikes on employment transitions in Brazil. We will estimate the probability of becoming nonemployed (unemployed or out of the labour force) and the probability of moving from the formal to the informal sector after minimum wage hikes[3]. We will estimate these effects separately for periods with high and low inflation to assess how agents react to minimum wage hikes under different inflationary expectations, particularly, under different degrees of wage indexation. Workers affected by minimum wage increases are compared with similar workers further up the wage distribution. In

[3]We estimate these effects for the joint sample of formal and informal sector workers and separately for each sector.

order to control for heterogeneity between the treated minimum wage workers and the comparison groups we use a difference-in-differences approach that compares treated and comparison groups in periods with nominal increases in the minimum wage with periods with no increase. In this last case the comparison and treated groups are defined as if there had been an increase in the minimum wage (pseudo-experiment). This strategy is applied in a parametric way via probit estimates for a variety of comparison groups and also in a nonparametric way using different propensity score matching methods.

The main findings are that disemployment effects are much more likely to be found in the late 1990's (period without official indexation) than in the early 1980's. We do not find strong evidences of transitions from the formal sector to either informality or self-employment after minimum wage hikes. Methodologically, probit and propensity score matching difference-in-difference approaches do not differ much in their estimates, mainly in relation to kernel estimates.

5.2 The Recent Literature on the Disemployment Effects of the Minimum Wage

5.2.1 The International Evidence

The disemployment effect of increases in the minimum wage is one of the main arguments against its widespread use as a redistributive policy. Card and Krueger's (1995) influential book *"Myth and Measurement: the new economics of minimum wage"* shed some doubts on the conventional wisdom that minimum wage hikes would trigger disemployment effects whenever the minimum wage is binding. The time-series based consensus of a 1 to 3% disemployment effect of a 10% increase in the minimum wage established in Brown et al. (1982) was seriously challenged. Applying several different methodologies and emphasizing (quasi)natural experiments, Card and Krueger provoked a revival in the empirical minimum wage research in both the US and worldwide, and also led researchers to look for economic models that would explain small negative effects or even positive effects of the minimum wage. Monopsony-like models and monopolistic competition are the basis of most of these models where small increases in the minimum wage do not hurt employment and in some cases can even increase it[4]. However, the main effect of the book was concentrated on the empirical research and in the use of several techniques that

[4]See, for instance, Card and Krueger (1995, ch.7), Rebitzer and Taylor (1991), Bhaskar and To (1996) and Brown (1999) for a review.

have not previously been applied in this field[5] and that aimed to uncover the real (causal) effect of minimum wage on employment-related outcomes. Recent findings of no disemployment effect or even small positive employment gains, such as the ones reported in Zavodny (2000) for the US, Portugal and Cardoso (2001) for Portugal, and Stewart (2002) for the UK have given support to the Card and Krueger challenge. Brown (1999) in his survey on minimum wage research concludes that, all in all, the recent evidence shows that the minimum-wage effect is small[6] and that zero effect is often hard to reject. At the same time, however, there is no shortage of recent papers that find negative effects of minimum wage on employment as predicted by the standard competitive labour demand model. Abowd et al.(2000) and Kramarz and Phillipon (2000) find that real increases in minimum wage led to higher probability of transition to unemployment in France. Pereira (1999) finds that minimum wage hurt teenage employment in Portugal, and Machin et al. (2002) find evidence of employment and hours reduction for care workers in the UK after the introduction of a national minimum wage in 1999[7].

The microeconometric literature on the disemployment effect of the minimum wage has focused on developed countries. In the US literature most papers estimate the minimum wage effect on teenage employment, which is the group most likely to be affected by the minimum wage hikes. Methodologically these studies try to use unaffected control (or comparison) groups to assess the "causal" impact of minimum wage hikes on affected workers. Card and Krueger (1995) use different minimum wage policies followed by similar states in the US to measure the effect of the minimum wage on employment - a quasi-experimental strategy. Pereira (2001) uses the increase in the minimum wage for teenager workers in Portugal as a quasi-experiment where older workers - not affected by the increase - are used as control group. Zavodny(2000), Kramarz and Phillipon (2000) and Abowd et al.(2000) use workers further up the wage distribution as a control group for workers affected by minimum wage hikes. To address the bias caused by time-invariant unobserved

[5] As a matter of fact, most of these techniques were used in the past, but lacked good data and good comparison groups. These difficulties were overcome in most recent studies by tailor-made surveys to evaluate the impact of the minimum wage (e.g. the fast food study in the US by Card and Krueger (1995) and home care in the UK by Machin et al. (2002)) and by the increase in the number of state minimum wage above the federal one in the US case, making possible to use regional variation to assess the impact of the minimum wage hikes. Also contributed to the increase in the number of empirical studies the fact that in many other countries household survey microdata became publicly available.

[6] Brown points out that even the more recent time-series studies for the US point to a lower than 1% effect (and in general, not significant) when the 1980's data are included in the estimation.

[7] However, the authors point out that their results cannot be extended to the whole economy, since it is a case study of a sector where the minimum wage is most likely to bite.

heterogeneity[8] between minimum wage workers and workers further up in the wage distribution, they build pseudo-experiments in a differences-in-differences context in order to have a baseline to compare the effects of real increases in the minimum wage for treated and comparison groups. The pseudo-experiment consists in pretending that there had been a minimum wage increase in a year[9] when it did not happened. Then, workers whose wages are between the current and the pseudo new minimum wage are defined as the treated group and workers marginally above as the comparison group. In this approach, the second difference should guarantee that the remaining difference in outcomes - e.g. in wage, hours and employment transition - between treated (workers between the old and the new minimum wage) and control groups can be more reliably identified as the "causal" effect of the minimum wage than either the results that one gets from the simple first difference between treated and comparison group in a cross-section or the simple first difference between treated individuals from a before-after analysis.

Using this methodology, Zavodny (2000) finds that the negative effect of minimum wage hikes on low-wage teens in the US disappears when the heterogeneity between this treated group and high-wage teens (comparison group) is controlled for using periods when the minimum wage did not increase. However, Kramarz and Philipon (2000) and Abowd et al. (2000) still find disemployment effect due to real increases in the minimum wage in France, even when controlling for heterogeneity via pseudo-experiment.

5.2.2 Evidence for Brazil

Lemos (2002) using a panel for six metropolitan areas in Brazil for the period 1982-2000 shows through kernel densities and descriptive regression analysis that increases in the minimum wage compress the earnings distribution (pooling data for both formal and informal sector workers) and that this effect is stronger in poorer than in richer regions. She also shows that the negative impact of minimum wage on employment is small and occurs mainly through a reduction in hours worked rather than in the level of employment. This small negative effect is observed in both formal and informal sectors. She argues that as wages increase with the minimum wage hikes, employment losses in both sectors should be observed.

Carneiro (2000) using time-series cointegration analysis for the period 1982-1999

[8]This unobserved heterogeneity may be what makes minimum wage workers more likely to become nonemployed even in the absence of a minimum wage hike than other workers.

[9]Most papers look at yearly transitions, in the case of Brazil, we will look at monthly transitions due to the high turnover rate and frequent wage hikes.

finds that changes in the minimum wage have robust and negative impact on the employment level of formal sector workers in the long run and positive impact on the employment levels of informal sector workers[10]. He argues that minimum wage hikes curb formal employment and forces dismissed workers to migrate to inthe formal sector, however the absolute value of the estimated employment-elasticities are quite low[11]. It is worth noting that his finding is in line with Welch (1974) model. Carneiro also finds that the short run wage-elasticity in relation to the minimum wage of the informal sector (0.24) was higher than the wage-elasticity of the formal sector (0.10), suggesting that in the short run, increases in the minimum wage have a major impact on average wage of the informal sector.

Neri (1997) points out that the bulk of poverty reduction witnessed after the 1994 *Plano Real* was observed just after the 42.5% nominal increase of the minimum wage in May 1995 and not immediately after the stabilization plan[12]. He also points out that formal sector workers between the old and the new minimum wage of May 1995 have higher (double) probability to migrate to the informal sector than the group marginally above, and their probability of becoming unemployed or moving out of the labor force is also higher[13]. Another interesting finding reported in Neri's paper is that the percentage of informal sector workers getting wage increases equal to the increase in the minimum wage was 21.5%, much higher than the 12% of formal sector workers in the same situation. In fact the proportion of informal sector workers with wage hike equal to the minimum wage increase has been increasing steadily since 1990[14]. During the 1980's, however, only a tiny proportion[15] of informal sector workers had their wages pegged to the minimum wage increase[16].

The high proportion of informal sector workers with wage increases equal to the minimum wage hike led Corseuil and Morgado (2002) to disregard this group as a good control group in their estimation of the impact of minimum wage hikes on the employment of workers earning between the old and the new minimum wage. In their evaluation of the minimum wage hikes between 1995 and 1999 for formal and

[10]The estimated elasticity for formal sector workers varies from -0.001 to -0.024 and for informal sector workers from 0.0004 to 0.003.

[11]Notice that this conclusion is at odds with Lemos (2002) results.

[12]Foguel et al. (2000) estimate that a 10% increase in the minimum wage would reduce poverty by 4% based on the minimum wage hikes from 1995 to 1998.

[13]Notice however that such evidence is based on raw probability transitions. It does not control for observable and unobservable heterogeneity between treated and control groups, neither tests the statistical significance of the differences.

[14]See more evidence on that in subsection 5.3.2

[15]This proportion was much lower than the one observed for formal workers.

[16]Neri (1997) also points out that after the stabilization plan the proportion of workers with wage increases equal to the increase in the minimum wage increased considerably - from 6% in average between May/80 - February/86 to 14% between September/94 to May/95.

informal workers separately, they did not find any consistent pattern on the effect of the minimum wage on the employment probabilities of affected workers[17]. They estimate a difference-in-difference logit model for transitions out of employment just after the minimum wage increase and the transition observed for the pair of months just before the increase in the minimum wage is used as counterfactual.

5.3 Minimum Wages in Brazil

The minimum wage was introduced in Brazil in 1940 with the aim of fighting poverty. However, its coverage has never been complete since informal sector employers do not necessarily comply with the labour legislation in general, and with the minimum wage policy in particular. The historical trend of the real value of minimum wage followed a cycle of political changes, experiencing increases during populist and left-wing leaning governments and being squeezed during conservative governments, when it was seen as an instrument to control aggregate demand in order to fight inflationary wage-price spirals.

5.3.1 The Minimum Wage Policy between 1982 and 1999

Since 1965 minimum wage hikes were linked to the wage policy implemented by the military dictatorship. The idea was to break the link between nominal wage increases and past inflation, readjusting wages by their average real value of the last 24 months plus a productivity index. The implementation of this policy was made easier by the nature of the regime and by the intervention in the trade unions. This policy was successively modified via the reduction of the months taken into account for the calculation of the wage increase. By 1979 the policy had changed with the introduction of wage increases every 6 months and with the introduction of different formulas for different ranges of the wage. The minimum wage and the range of the wage up to 3 minimum wages should be increased by 1.1 of the accumulated inflation of the past 6 months. For the range between 3 and to 10 minimum wages the index was 1, and above this, 0.8. In 1983, the government ended the over-indexation to inflation of the minimum wage and of the wage range up to three minimum wages, reducing the index from 1.1 to 1 and decreased the index for the range between 3 and 10 minimum wage from 1 to 0.8. Had this policy been successful, it would have

[17]Workers earning up to 2 minimum wages were used as control group, excluding the ones who earned exactly 1.5 minimum wage or 2 minimum wage. It seems that the statistical significance of most of their estimates was due to the use of frequency weights in the estimate which decreased a lot the standard error of the estimates.

implied a compression in the wage distribution, something that was not observed. The minimum wage during this period worked as an index for several prices in the economy and as the base wage rate for several unionized workers in different sectors of the economy.

In 1985 the military rule came to an end, and the new civilian government tried to increase the real value of the minimum wage by conceding increases above the past inflation rate. In February 1986 the government announced the *Plano Cruzado* and prices and wages were frozen. Wages were frozen by their mean real value in the past six months plus an 8% bonus. For the minimum wage this bonus was 16%. According to the new wage policy wages would be increased automatically when the accumulated inflation reached 20%.

The *Plano Cruzado* failed to control inflation and after its failure, successive attempts to halt the hyper-inflationary process either via restrictive fiscal and monetary policies or wage and price controls were tried, but none succeeded. Wages alternated between periods of monthly increases and some "longer periods" of wait for readjustment, but never less than six months. Nevertheless, this climate of uncertainty and rising inflation led to a fall in the real value of wages, in general, and also of the minimum wage. An additional factor that marked the minimum wage policy during this period was the enactment of the 1988 Constitution that linked several social benefits and pensions to the minimum wage. Due to this indexation, minimum wage hikes became a major issue for the fiscal balance of the government. Each increase in the minimum wage would generate pressure on the budget[18]. Thus, when deciding by how much to increase the minimum wage, the government did not only take into account its impact on the economy via employment and wage effects, but also its effect on the budget[19].

In 1994, the stabilization program known as *Plano Real* indexed prices and wages to the unit of real value (URV) and provoked a sort of controlled hyper-inflationary process that ended up with the change of the currency from *cruzeiro real* to *real*. Among several measures to avoid the new currency being contaminated by past inflation of the old currency, the government promulgated a law that forbade the indexation of wages to the inflation or to the minimum wage. In this context, the change in legislation brought about in 1994 represented a significant move towards the end of wage policy based on past inflation indexation.

[18]Another source of pressure on the budget is a high proportion of state and municipality public servants who are minimum wage workers.

[19]The government tried several times to break this link without success.

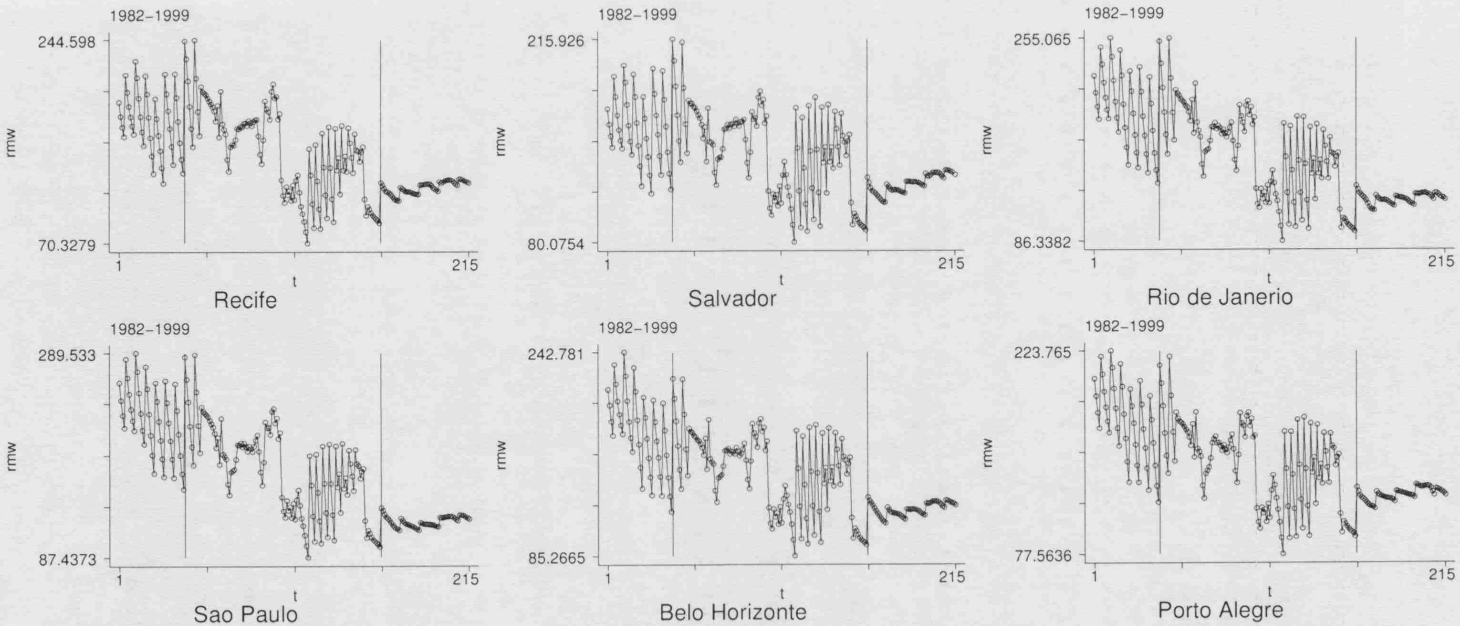

Figure 5.1: Real Minimum Wage in Six Metropolitan Areas - 1982 - 1999

The minimum wage was increased by 42.5% in May 1995 (in nominal terms) and since then has been increased every year in that month[20]. Figure 5.1[21] depicts the evolution of the real minimum wage for the 6 metropolitan areas for which we have monthly data[22]. It is striking that only after 1995 (the period on the right of the second bar) the graph resembles what one would expected from a real minimum wage evolution with a regular saw-toothed pattern. The period 1982-1985 also shows a similar pattern but with a smaller interval between the spikes (the period on the left of the first bar). Due to the irregularities of the changes in the minimum wage policy and the difficulty in getting clean data because of the extreme high inflation between 1987 and 1994, we decided to concentrate our analysis in two sub-periods: 1982 - 1985 and 1995-1999. These two periods are the ones with longer spells between minimum wage hikes, so that we can more easily find counterfactual months without minimum wage increases. Nevertheless, these two periods differ in relation to the way other wages were readjusted. Whereas between 1982-1985 there was legal indexation to past inflation, between 1995-1999 this indexation was forbidden[23]. Moreover, several institutional changes occurred between these periods, among them we can point out the 1988 Constitution and the increase in labour costs that it brought about and the trade liberalisation of the early 1990's[24]. In the labour market, one of the most striking changes was the increase in the proportion of workers in the informal sector (without registration).

5.3.2 A Tale of Two Decades and Two Sectors

As mentioned in the introduction of this chapter, the role of the minimum wage in Brazil as a reference for wage in the formal sector is a major characteristic of the distribution of wages in Brazil due to the wage policy of the 1980's. However, its role as an index was not restricted to the formal sector. In this regard, there is clearly

[20]The only exception was in 2000 when the minimum wage was increased in April.

[21]the horizontal axis varies from 1 to 215, the number of months between April 1982 and November 1999.

[22]We display the graphic for the 6 metropolitan area for two reasons. First, the minimum wage for Salvador and Recife was lower than for the other metropolitan areas between 1982 and 1983, when there was regional variation in the minimum wage, and second we used the Consumer Price Index (INPC) for each area in order to convert the nominal values into real values.

[23]Corseuil and Morgado (2002) argue that only for the period 1995-1999 it is possible to create credible control groups further up the wage distribution due to the difficulty to disentangle the effect of the minimum wage in the whole wage distribution before 1994.

[24]Chapter 4 of this thesis and Goldberg and Pavcnik (2003) show that trade liberalization cannot explain the increase in the proportion of non-registered workers. In this same vein, Gonzaga et al. (2002b) show that the 1988 reduction in the legal workweek also does not explain the increase in the proportion of non-registered workers.

two different phases. One between 1982 and 1985 when the minimum wage was an index for the formal sector, and had little - if any - impact on the wage distribution of the informal sector, and the period 1995 - 1999, when the minimum wage seems to be much more important for the determination of the wages in the informal sector than in the formal sector. One possible explanation for this phenomenon was the fall in the real value of the minimum wage observed during the 1980's and the first half of the 1990's as depicted in Figure 5.1.

To illustrate these changes we will analyse the kernel densities of the log wages in May 1985 and in May 1995 and also the wage changes by percentiles in periods of minimum wage increases. These two episodical minimum wage hikes are representative of the effects of minimum wage increase on the wage distribution for the early 1980's and for the late 1990's, respectively[25]. Figure 5.2 and Figure 5.3 show the distribution of log wages in the month before (April) and in the month after (May) a minimum wage increase for formal and informal sector workers in 1985 and 1995. The data come from the PME (Monthly Employment Survey) for the six main metropolitan areas.[26] The three bars in the graphs correspond (from left to right) to the minimum wage, to two times the minimum wage and to three times the minimum wage. Figure 5.2 reveals that the minimum wage produces a spike at its value in the distribution of wages for formal sector workers, and a second less pronounced mass at the value of two times its value. Moreover, there was a clear shift to the right from one month to the other after the minimum wage increase. The graphs in the second row of this figure show that this sort of truncation did not exist in the informal sector due to the high degree of non-compliance. More than a half of the informal sector workers earned less than the minimum wage in 1985. Nonetheless there is a spike at the value of the minimum wage, but not at its multiples.

Figure 5.3 differs from Figure 5.2 in several ways. First, non-compliance dropped considerably among both formal and informal sector workers between 1985 and 1995. Second, the most important spike in the wage distribution of formal sector workers occurs at two times the value of the minimum wage in 1995. There is still a spike at the minimum wage but it is less pronounced, particularly, in April 1995. Surprisingly, the graphs for the informal sector in 1995 resembles the ones for the formal sector in 1985. There is a spike at the minimum wage and a second spike at

[25]Whenever there is a significant difference between the results we get for these two specific years and the other years that they are meant to represent, we will comment on the differences. However, most of the stylised facts found for these two years also apply for similar years in the 1980's and in the 1990's.

[26]See section 5.5 for more details on this data set.

the value of two times the minimum wage. Moreover, there was a clear shift to the right of the distribution after the minimum wage increase. Therefore, the minimum wage seems to have affected the distribution of wages in both formal and informal sectors. However, during the late 1990's, it seems that the minimum wage became much more important for the distribution of wage in the informal than in the formal sector.

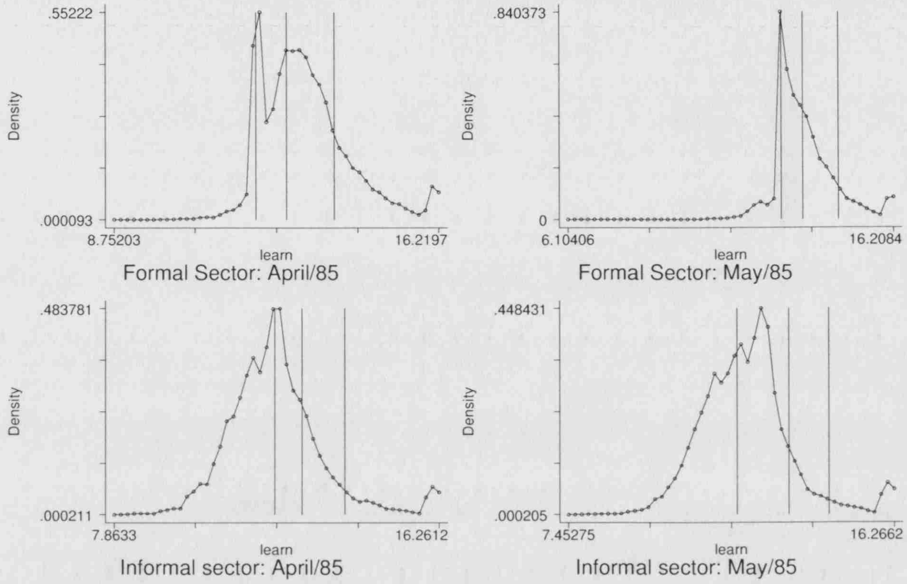

Figure 5.2: Kernel Density of Log Wages: 1985

In order to analyze the impact of minimum wage hikes on the entire wage distribution, we plot in Figure 5.4 the monthly increases in the average wage for each percentile (*diff1*) and compare it to the nominal minimum wage increase in May 1985 and May 1995 (the upper horizontal line in the graphs) for formal and informal workers, separately. In the case of absence of spillover effect, we expect to see spikes equivalent to the minimum wage hike for percentiles whose average wage are close to the minimum wage and no effect for highest percentiles. Due to the possibility that we would be capturing spurious increases in wages we also apply a differences-in-differences approach that controls for the variation in months surrounding the minimum wage increase: *diff2* and *diff3*. For 1985, *diff2* is the difference between the increase in the mean of the percentile in the month after the minimum wage hike (June) and the month before the minimum wage hike (April), and *diff3* is the difference between the increase in the mean of the percentile in the month after the minimum wage increase (May) and the average mean increase one month before (April) and one month after the minimum wage hike (June). For the

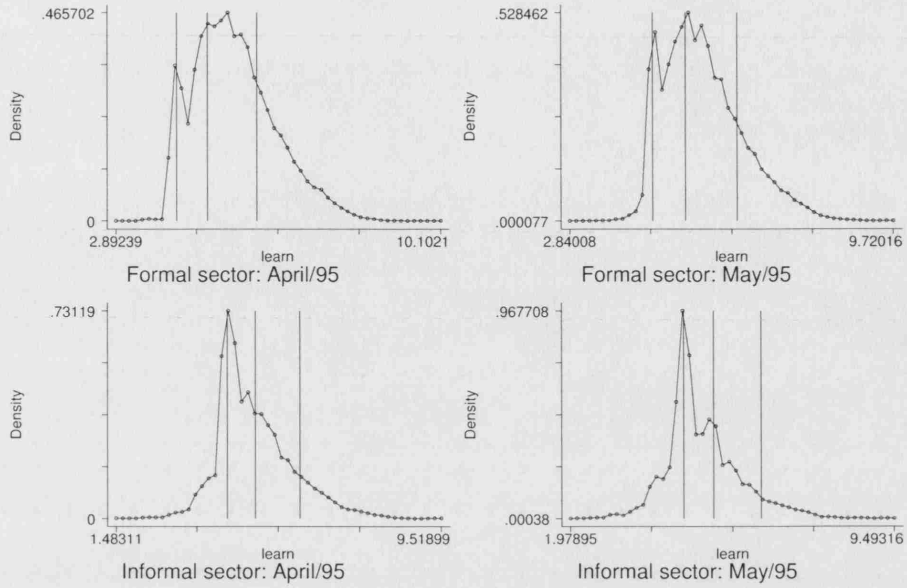

Figure 5.3: Kernel Density of Log Wages: 1995

period 1995, *diff2* is the difference between the increase in the mean of the percentile in the month after the minimum wage increase (May) and the average mean increase one month before(April) and one month after (June) the minimum wage hike[27], and *diff3* is the difference in the increase in the mean of the percentile in the month after the minimum wage increase and five months later (October)[28].

From the graphs in Figure 5.4 we can observe many contrasts between the percentile wage variation for formal and informal workers. First, the minimum wage hike in May 1985 had no impact on wage increases in the informal sector. The percentiles close to the minimum wage (first vertical line from left to right) had an increase slightly above 50% - the highest increase observed for all percentiles -, whereas the minimum wage increased by 100% (in nominal terms). Second, formal sector workers in the lower percentiles of the distribution had wage increases equal to the increase of the minimum wage (this can be seen by the flat upper line in the graphs). Notice that in general this flat portion of the graph coincides with the vertical line on the left that indicates the first percentile whose average wage is equal to the minimum wage. This coincidence in nominal increases occurred in both periods analysed. Third, in May 1995 many percentiles at the lower end of the wage distribution for informal sector workers had wage increases equal to the minimum

[27]Note the definition of *diff2* for 1995 corresponds to the definition of *diff3* for 1985.

[28]As there was an increase in the minimum wage in November 1985, we cannot apply the control group *diff3* as defined for the year 1995, i.e, using months well ahead the May minimum wage hike, as a control group for the year 1985.

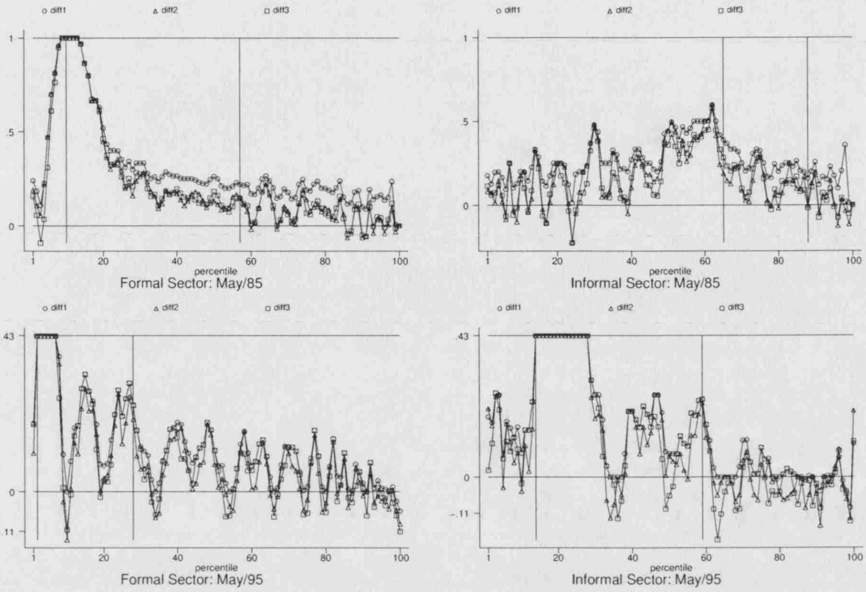

Figure 5.4: Wage Change by Percentiles due to Minimum Wage Increase: 05/1985 and 05/1995

wage. This represent a completely change in relation to May 1985. Actually, more percentiles in the wage distribution for informal sector workers had wage increases equal to the increase in the minimum wage than in the formal sector[29].

At a first glance, the graphs suggest that spillover effects were more likely to happen in 1995 than in 1985[30]. However, this only means that in the month of a minimum wage increase during the 1980's only the percentiles directly affect by the minimum wage seems to show an expressive positive variation. Nothing prevents that the other percentiles catch up with the minimum wage increase over time. Nevertheless, this indicates that during the 1980's every minimum wage increase in the very short run implied that the relative wage of minimum wage workers did increase.

As for indexation of multiples of the minimum wage for both minimum wage hike episodes reported in the graphs, May 1985 and May 1995, there is no clear spike at two times the value of the minimum wage. However, graphs not shown here reveal that from 1996 onwards there is a spike equal to the increase in the minimum wage near the percentile correspondent to two times the minimum wage. Thus, it seems that multiples of the minimum wage in the late 1990's were much more indexed to

[29] This phenomenon was first reported in Neri(1997) as seen in the last section.

[30] Actually the figures for minimum wage increases from 1995 onwards show even stronger spillover effects.

it than in the 1980's.

5.4 Methodology

In this section we will discuss the methodology applied to estimate the effect of the minimum wage on employment transitions for the pooled sample of formal and informal workers and for each sample, separately. We will apply difference-in-differences with pseudo experiment using two different techniques - probit estimates and propensity score matching - and different comparison groups.

5.4.1 Difference-in-Differences with Pseudo-Experiment

The only difference between the traditional difference-in-difference approach and one with a pseudo-experiment is that in the standard case, the treated group is defined regardless of the existence of the treatment or not. For instance, union workers can be defined as the treated group in the case of an evaluation of a major union-rights reform: both unionized and non-unionized workers (comparison group) are clearly observed and defined before and after the treatment without any trick. This does not occur in the case of an evaluation of the effects of a minimum wage increase. In this case, one has to pretend that there had been a minimum wage increase in the period used as baseline (before or some time after the minimum wage increase) in order to define the treated group - workers between two different minimum wages - and the correspondent comparison group - workers marginally further up in the wage distribution. The (pseudo)increase used to define the treated and control group is arbitrary, which means that one can use different "(pseudo) minimum wage increases" in order to generate the counterfactual for the "actual minimum wage increase".

Figure 5.5 presents one example of how we define treated and comparison group in a "pseudo experiment" framework. It shows the wage distribution for the pooled sample of formal and informal workers for March 1995, April 1995 and September 1995. The April 1995 distribution is the one that was actually affected by the minimum wage increase of May 1995. The treated (affected) group is defined by the area between the first and second vertical bars from left to right, i.e, the area between the current minimum wage (April 1995) and the minimum wage for May 1995. The comparison group is defined as workers between the May 1995 minimum wage and two times the current (April 1995) minimum wage (the area between the second and third bars). The two other distributions will generate the counterfactual used to assess the impact of the minimum wage on the distribution of April 1995. For March 1995, the treated and comparison groups are defined in the same way that they had been defined for April 1995. Thus, we pretend that the May 1995 increase would have occurred in April 1995, so that treated and comparison groups can be defined based on the wage distribution of March 1995. For September 1995 the construction of the counterfactual treated and control groups is a bit different. In this case we pretend that the minimum wage hike that would occur in the following year (May 1996) was anticipated to October 1995, so that we can define treated and control groups using the wage distribution of September 1995. The treated group is defined as the area between the current minimum wage, which is the same as in May 1995, and the forthcoming one (May 1996), and the comparison group is defined by the area between the forthcoming minimum wage and two times the current one[31]. The actual effect of the May 1995 minimum wage hike is estimated by comparing the estimated differences in transitions for treated and control groups for the pair April/May 1995 with these same differences calculated for both counterfactual pairs: March/April 1995 and September/October 1995.

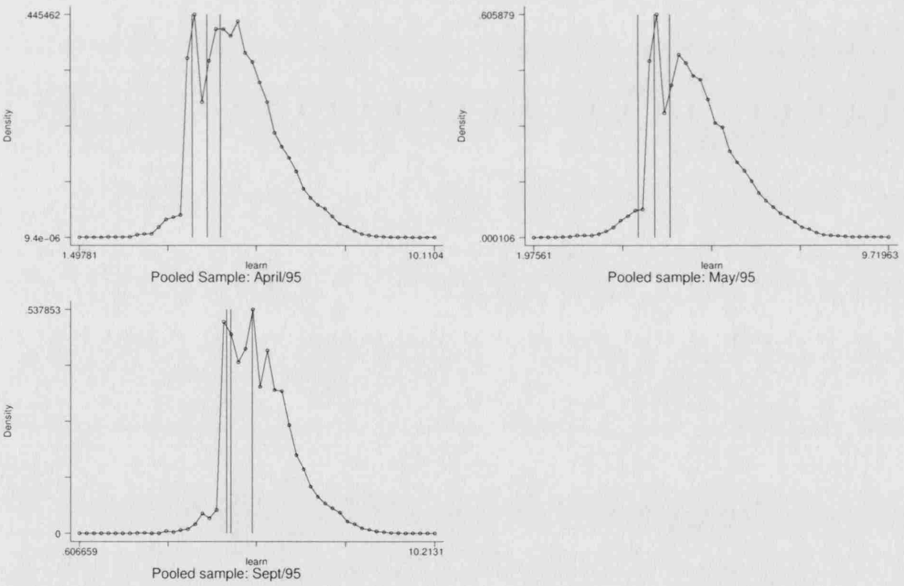

Figure 5.5: Wage Distribution and Treated and Control Groups: 04/1995, 05/1995 and 09/1995

A common characteristic of the strategies based on the idea of differences-in-

[31]Actually, these are the two counterfactuals that we are going to implement for the subperiod 1995-1999. We will compare changes in transitions in months of minimum wage increase with the transition immediately before the minimum wage increase and five months later. For the period 1982-1985 we will use only the transition immediately before the minimum wage hike.

differences with pseudo-experiment is that they rely on parametric models - mostly linear probability models, logit or probit analysis - in order to measure the impact of minimum wage hikes on both treated and comparison groups. The caveats of these parametric models in dealing with comparisons between treated and comparison groups are well-known. First there is the problem of the lack of common support for treated and comparison groups. This feature implies that the estimation of the effect of the minimum wage relies on linear (or other assumed functional form) extrapolations to evaluate the impact over regions where there is no available counterfactual in the comparison group for the treated group[32]. Second, in a difference-in-difference context, changes in group composition over time may lead to an imbalance between the observed characteristics of both comparison and treated groups after the treatment rendering the comparison meaningless[33]. Due to these shortcomings, this paper also applies propensity score matching methods to evaluate the impact of the minimum wage on employment transitions. In doing so, we will be building comparison groups that are more reliable than the ones available for the parametric analysis[34].

5.4.2 Choosing control groups

Assuming that Y_{it}^1 is the employment status for the individual i in the treated group[35] in period t after a minimum wage increase, the effect of the minimum wage increase on this individual is given by $Y_{it}^1 - Y_{it}^0$ and the average effect of the treatment on the treated (ATT) can be defined as $E(Y_{it}^1 - Y_{it}^0|I = 1)$. The problem with this approach is that there is a missing data: $E(Y_{it}^0|I = 1)$ cannot be calculated since Y_{it}^0 is not observed for treated individuals. The two treatment states are exclusive in period t. We do not observe the outcome of non-treatment for someone who was treated, in our case, we do not observe the employment status of a treated individual if the minimum wage hike did not happen. Thus the task is to find a comparison group, not submitted to the treatment, i.e., not affected by the minimum wage hike

[32]See Smith and Todd (2000) for an in-depth discussion of this issue.

[33]See Blundell et al. (2001) for the importance of controlling for compositional effects in a difference-in-differences context. However, this difficulty can be overcome in a parametric framework by using changes in the characteristics rather than the characteristics themselves as explanatory variables in case one has access to a panel.

[34]Brown (1999) argues that most of the early studies that applied difference-in-differences to the evaluate minimum wage effect had built naive comparison groups. In fact, when using workers further up the wage distribution the reliability of the comparison group is even more difficult to prove due to possible existence of both substitution and spill-over effects.

[35]Treated group is defined as workers between the old and the new minimum wage - $[I(mw_{t-1} <= W_{it-1} < mw_t) = 1]$. The superscript 1 in Y_{it}^1 indicates the treated status. Similarly, 0 in Y_{it}^0 indicates the absence of treatment.

- $(Y_{it}^0|I = 1, mw = 1) = (Y_{it}^0|I = 1)$ where *mw=1* indicates the minimum wage hike) - whose characteristics are similar to the affected individuals so that we can approximate Y_{it}^0 for the group of treated individuals

If one could find a comparison group with similar observable and unobservable characteristics and whose outcome follows a similar trend as the one observed by the treated group before the treatment, then the difference-in-differences estimator would be the natural candidate to recover the causal effect of the treatment - a minimum wage hike - on the treated:

$$\alpha = E(Y_{it}^1 - Y_{it}^0) - E(Y_{it'}^1 - Y_{it'}^0) \tag{5.1}$$

where t' is some period of time before or after the treatment when the treatment effect is assumed to have ceased[36].

The advantage of using a difference-in-differences approach can be better seen if we analyze the shortcoming of the two components of the difference-in-difference technique. For instance, the simple difference between treated and control groups after treatment, i.e., after the minimum wage hike, can be due to the remaining unobserved heterogeneity between the two groups and not to an actual effect of the treatment[37]:

$$E(Y_{it}^1 - Y_{it}^0) = \alpha + (\gamma_1 - \gamma_0) \tag{5.2}$$

where $\gamma_j, j = 0, 1$ is the time invariant heterogeneity for each group. One way to get rid of this individual specific heterogeneity is to take the difference between treated individuals before and after the treatment, the so-called *before-after estimator*. The shortcoming in this case is that there is no way of disentangling the actual effect of the treatment α from any other macro shock that would have altered the trend of the outcome variable for treated individuals:

$$E(Y_{it}^1 - Y_{it'}^1) = \alpha + (\phi_t - \phi_{t'}) \tag{5.3}$$

where $\phi_j, j = t, t'$ is a common macroeconomic effect. Thus, only if we take the difference between the outcome for the two groups in those two time periods - as in equation (5.1) - we would get the "causal" effect of the treatment purged of both unobserved time-invariant heterogeneity, γ_j, and common time trend, ϕ_j. The importance of an appropriate comparison group in this case is to cancel out any

[36]In general t' is defined as $t' < t$, but we also want to allow a "control period" after the treatment, assuming that the minimum wage hike only affects employment in the short run.

[37]This is the typical bias found in cross-section estimation.

common time effects that would bias the estimator. Therefore one of the key assumptions of the difference-in-difference model is that both treated and comparison group must display the same - or as similar as possible - time trend before the treatment and that no shocks occurred simultaneously with the minimum wage hike that could differently affect the control group and the treated group[38].

In a difference-in-differences framework both unobserved time-invariant heterogeneity and common time effects between treated and comparison groups could be eliminated by taking the differences between them in two different time periods, one period under the effect of the treatment and another period when the treatment did not occur, but when appropriate treated and control groups can be defined (pseudo-experiment), and then taking the difference of these differences.

The common trend assumption implies that changes in the outcome in the absence of a minimum wage hike for treated workers should be the same as the change observed for workers in the control group:

$$E(Y_{it} - Y_{it'}) = [E(Y_{it}^0 | I = 1) - E(Y_{it'}^0 | I = 1)] = [E(Y_{it}^0 | I = 0) - E(Y_{it'}^0 | I = 0)] \quad (5.4)$$

In the perfect competition framework, the traditional theory of labour demand assumes that an increase in a binding minimum wage renders workers whose productivity is below the newly set value unemployable. Thus minimum wage increases would tend to increase non-employment either through increases in inactivity or in unemployment. A potential problem for the construction of good comparison groups for minimum wage workers is the fact that workers marginally further up the wage distribution can be close substitutes for minimum wage workers so that their employment probability can be (positively) affected by an increase in the minimum wage, depending on the elasticity of substitution between the two groups and on the magnitude of the increase.

If this is the case, the outcome trend before and after the treatment cannot be the same for both groups and (negative) minimum wage effects (on the treated) would be overestimated. The assumption here is that the difference in the outcome between the treated and the comparison groups should be constant in the absence of the treatment, so that the comparison group would offer a good approximation of what would have been the outcome for the treated group in the absence of the

[38] Blundell and Dias (2000) describe the differential-trend-adjusted difference-in-differences estimator as way to overcome this problem. The idea is to find a similar time interval in the business cycle where the treatment did not happened and calculate a counterfactual difference-in-difference estimator that would be subtracted from the difference-in-difference estimator over the time interval when the treatment did take place.

experiment - an increase in the minimum wage.

Kramarz and Phillipon (2000) point out that the demand for workers in the comparison group must not vary with changes in the minimum wage. One alternative control group would be workers who earn well above minimum wage workers, but in this case workers in the control group would not be comparable to minimum wage earners since their observable and unobservable characteristics would be quite different, making even worse the problem of lack of common support[39]. Testing whether or not comparison groups are affected by minimum wage hikes is important in the sense that depending on the elasticity of substitution between treated and comparison groups used in the analysis, we can be over- or underestimating the effect of the minimum wage. Kramarz and Phillipon (2000) put forward a test of the quality of the comparison group through the difference in elasticities of employment probability during minimum wage hikes and during pseudo-minimum wage hikes - the difference between the elasticities calculated in these two periods should be small and statistically insignificant for the control group to be considered a good one. This is particularly important in the case of Brazil, since as pointed out in the last section, there has been a history of wage indexation in relation to both inflation rate and minimum wage increases.[40].Figure 5.6 show the job loss pattern throughout the wage distribution in the month before (April) and in the month after (May) the minimum wage increase. It shows that job losses are much more prevalent for the lowest percentiles and that the job loss rate in months following minimum wage increases is slightly higher than in months immediately before for this group of workers. As minimum wage earners are concentrated in the lowest percentiles (in 1995 they were around the 10th percentile), choosing a comparison group well above the minimum wage can potentially lead to an upward bias of its negative effect, since further up the wage distribution the job loss pattern is very different.

Other alternatives to the use of workers further up the wage distribution as comparison groups would be to exploit regional variation in minimum wage hikes[41] or the use of workers in the informal sector as an alternative comparison group.

[39]This problem is quite common in approaches that contrast the probability of employment in t for individuals that in $t-1$ earned between the old and the new minimum wage with the probability of employment of all other workers, the general procedure to overcome the criticism of bad quality comparison group has been the inclusion of dummies for the group marginally above the minimum wage and testing whether the impact was different for this group and the treated one, see for instance Currie and Fallick (1996) and Abowd et. al (1997).

[40]As seen in last section, spillover effect does not seem to be a major problem in Brazilian data. Lemos (2002) and Soares (2002b)also find small spillover effects beyond the 10th percentile of the wage distribution.

[41]Foguel (1998) pursued this alternative.

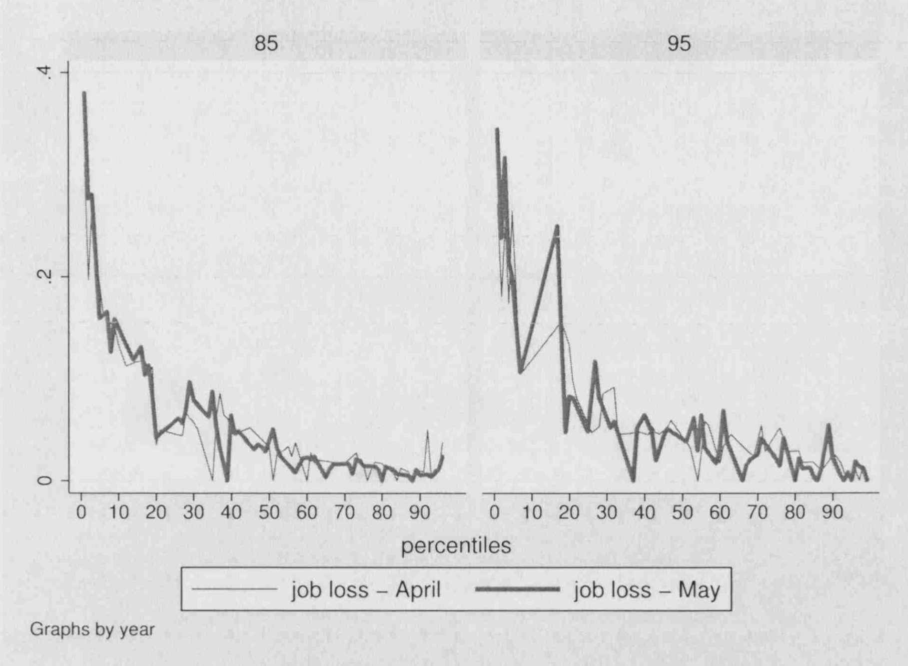

Figure 5.6: Job loss before and after minimum wage increases in 1985 and 1995

The problem with the first alternative is that it would be limited to the period between 1982 and 1984 when there was regional variation in the minimum wage in our data. The second alternative could be pursued using non-registered (informal) workers as a comparison group. However it also has its limitation because, as seen in the last section, there is evidence that non-registered workers wages are directly affected by minimum wage hikes, at least in the period after 1995. In this case their appropriateness as control group would be questionable given that they would be at risk of losing their job as any registered (formal) worker[42]. Actually, the possible direct effect of minimum wage hikes on informal sector employment makes it another possible treated group. For this reason our analysis of the disemployment effect of minimum wage hikes will be applied to all employees in both formal and informal sector, and separately for each group.

Our preferred control group will be workers further up the wage distribution up to the monetary value of two old nominal minimum wages[43] in periods when

[42]However, it is important to keep in mind that employers of informal workers are not constrained to comply with the minimum wage legislation as are formal employers. So it may be the case, that the effect of minimum wage on informal workers is much milder than the one would expect for formal workers

[43]The only exception is the 100% nominal increase observed in May 1985. If we use 2 minimum wages as the cut point we would end up with an empty comparison group. Thus for this particular event we use 3 nominal minimum wages as the upper bound of the control group.

minimum wage hikes effectively occurred. A similar definition will be applied in our pseudo-experiment dates, when we will pretend that there had been a minimum wage hike one month before the minimum wage increase and another five months after the increase. These pseudo-experiments will act as a counterfactual for the real experiment in order to control for unobserved and time-invariant heterogeneity between treated and control groups.

Due to the existence of a large number of workers whose wages are equal to multiples of the minimum wage we repeat the estimation above excluding them from the control sample[44]. It is interesting to observe that if this group really follows the minimum wage policy, minimum wage hikes should affect it in the same way it affects minimum wage earners, thus their inclusion in the control group should, if anything, lessen the effect of the minimum wage on the treated group.

In order to test the robustness of the results of our preferred comparison group we also test two other comparison groups in the probit specifications: workers earning more than two old minimum wages and workers earning less than the old minimum wage.

The aim of these checks is to build comparison groups that closely resemble minimum wage workers in observable dimensions, that are not affected by minimum wage hikes and whose remaining unobserved differences may be purged by taking the differences between them in two time periods. Given the clear impossibility of building a perfect comparison group in a non-experimental context, we try to overcome this difficulty using different comparison groups and assessing how closely they are to fulfil the assumptions necessary to validate the results.

5.4.3 The Common Support Problem and Matching Methods

In general, probit or logit models are used to estimate difference-in-differences coefficients for dichotomous outcome variables. The parameter of interest in this case is the coefficient of the interaction between a dummy that identifies the treated group and a dummy that identifies the period after the treatment as α in the equation below:

$$Y = \Phi(X_t\beta + \alpha I * t + \gamma I + \lambda t) \tag{5.5}$$

where Φ is the standard normal distribution, X includes observable characteristics

[44]Courseuil and Morgado(2002) use this control group for their logit analysis.

that affect the binary outcome Y, β is a vector of coefficients, I is the indicator for the treated group and t is the period indicator.

The problem with such specification is twofold. First, it does not take into account that some comparison group members may be a bad match for some treated individuals. Second, it overlooks the fact that the treatment may lead to changes in the X's of the treated group causing compositional changes in both groups.

This first problem is known as the failure of the common support, which forces the data to find a counterfactual in the comparison group for a treated individual even when in fact there is none available. Such comparisons are based on extrapolations according to the assumed functional form in order to provide the contrast. There is no guarantee that this extrapolation is accurate. Thus, if one is concerned about the average treatment effect on the treated, a non-parametric approach that would impose common support condition on both treated and comparison group would be more likely to offer a good guide of the actual effect of the treatment on the treated.

A popular non-parametric approach to tackle this problem in the programme evaluation literature is matching methods. The procedure consists of several techniques that offer for each treated individual a close match in the comparison group. This match can be simply an individual with similar observable characteristics, or the weighted or unweighted mean of a set of individuals with similar characteristics. The methods also differ according to the way that the mean is calculated and whether or not it imposes restriction on the possibility of matching. That is, whether or not it allows treated individuals not to be matched in the case where the controls are not very good according to some criteria. After a suitable match pair is found or built, it is possible to estimate the average of the two groups and then take the difference, the result is the average treatment on the treated.

A major problem with the earlier versions of matching methods was the question of dimensionality. Matching treated individuals with similar control individuals requires controlling for all relevant variables that determine whether an individual is in the treated group, so that any remaining bias in the allocation to the treated group would be due to unobservable characteristics. However, this procedure can make matching unfeasible as the number of cells in which dimensions (X) individuals should be matched would increase, demanding huge data sets rarely available to researchers. Moreover, in the case of discrete variables, one would end up with several null cells and in the case of continuous variables the rate of convergence of the estimator would be severely reduced (Smith and Todd, 2000).

This problem is overcome with the use of propensity score, $Pr(I = 1|X) =$

$P(X)$, instead of the X observable variables in the conditioning[45]. This procedure allows the reduction of the multidimensional problem, to a one-dimension (or to a lower order manageable) problem. By this method, only individuals with similar propensity score would be compared. The mean independence assumption necessary to recover the average treatment on treated (ATT) can be stated as:

$$E(Y_{it}^0|P(X), I = 1) = E(Y_{it}^0|P(X), I = 0) \tag{5.6}$$

In addition we also need to assume that all treated agents (at least within the common support) have a counterpart on the non-treated population and anyone included in the estimation constitutes a possible participant:

$$0 < P(I = 1|X) =< 1 \tag{5.7}$$

Thus the first step when applying matching is to estimate the propensity score associated to the treated group, which can be done parametrically via probit or logit estimation. The second step is to calculate the common support where the ATT can be defined, based on the overlap region of distribution of the propensity scores for both treated and control groups. Finally, the third step is to to build the estimator of the impact of the policy change:

$$\alpha_{mcs} = \frac{1}{n_1} \sum_{i \in I=1} [Y_i - \sum_{j \in I=0} W_{ij} Y_j] \tag{5.8}$$

where n_1 is the number of treated observations and W_{ij} is the weight associated to the controls ($I = 0$).

In order to do that, one has to define the number of control observations that are going to be used as contrast for the treated individuals and the weights, W_{ij}, associated with them. This is equivalent to the choice of the matching method to be employed. There are several options in the literature, but in this chapter, we are going to use two methods: the nearest neighbour estimator and the kernel estimator.

The Nearest Neighbour matching (NN) is the traditional pairwise matching. We match each individual in the treated group (inside the common support region) with the individual in the comparison group who is the closest in terms of their difference in propensity scores: $d(P_i) = min\|P_i - P_j\|, j \in I = 0$. It is possible also to use some minimum acceptance distance to avoid the occurrence of bad matches. As a matter of fact, the common support restriction is, in general, imposed *a priori*

[45]See Rosenbaum and Rubin (1983).

via this maximum acceptable distance[46]. The NN procedure may be extended via the use of more matches so that the counterfactual becomes the average of the number of matches used (each observation is given the same weight in the calculation of counterfactual). In this case, the variance of the estimator decreases, but this procedure can increase the bias of the estimate due to poorer matches. This chapter will report the results for the simple NN matching estimator:

$$\alpha_{nm} = \frac{1}{n_1} \sum_{i \in I=1} [Y_i - Y_j] \tag{5.9}$$

The kernel matching builds the counterfactual based on a kernel weighted average over a set of individuals in the control group:

$$\alpha_{km} = \frac{1}{n_1} \sum_{i \in I=1} Y_i - \frac{\sum_{j \in I=0} Y_j K(\frac{P_j - P_i}{h_n})}{\sum_{k \in I=0} K(\frac{P_k - P_i}{h_n})} \tag{5.10}$$

where $K(.)$ is a kernel function and h_n is a bandwidth parameter. The weight, W_{ij}, as in the general format of (5.8) corresponds to $\frac{K(\frac{P_j - P_i}{h_n})}{\sum_{k \in I=0} K(\frac{P_k - P_i}{h_n})}$ in the kernel matching. Heckman et al.(1997) show that under standard conditions on the bandwidth and the kernel[47], the counterfactual generated by this method is a consistent estimator of $E(Y^0|P(X), I = 1)$. The choice of the kernel and of its bandwidth will determine the number of control observations used to match each treated individual and the weight[48] with which each control observation will enter the calculation of the counterfactual. The Gaussian kernel, for instance, is unbounded which means that all control observations enter the calculation. In this chapter we use a Epanechnikov kernel with constant bandwidth chosen according to Silverman's rule of thumb.

Matching estimators are easily applied to a difference-in-difference framework as in equation (5.1) with minor changes in the way the cross-section matching is calculated. As in the parametric case, difference-in-differences has the advantage of eliminating any time-invariant unobserved heterogeneity that makes a certain individual more likely to be in the treated than in the comparison group. Besides, it also eliminates common trend effects that drive the outcomes. Its advantage over the parametric strategy of equation (5.5) is that it does not impose any specific functional form in estimating the conditional expectation of the outcome variables

[46]Notice that the same control can be used for different treated observations.

[47]These conditions require that $K(.)$ integrates to one, has mean zero and that $h_n \rightarrow 0$ as $n \rightarrow \infty$ and $nh_n \rightarrow \infty$.

[48]The weight of each control observation decreases with its distance from the treated observation as measured by their propensity scores.

and it reweights the observations according to the weighting functions used by the matching estimators. The assumptions needed to justify the application of this estimator are[49]:

$$E(Y^0_{it} - Y^0_{it'}|P(X), I = 1) = E(Y^0_{it} - Y^0_{it'}|P(X), I = 0) \qquad (5.11)$$

and $0 < P(X) < 1$, where t' indicates a period, in general, before the treatment, but that actually can be any period in which there is no impact of the treatment.

Assuming that we only have access to repeated cross-section, the estimator can be written as:

$$\alpha_{mdid} = \frac{1}{n_{1t}} \sum_{i \in I=1}^{n1t} [Y_{it} - \sum_{j \in I=0} W_{ij} Y_{jt}] - \frac{1}{n_{1t'}} \sum_{i \in I=1}^{n1t'} [Y_{it'} - \sum_{j \in I=0} W_{ij} Y_{jt'}] \qquad (5.12)$$

The problem here is how to determine the common support for the two periods of time under consideration, t and t'. One strategy would be to calculate the propensity score for the aggregate $P(I = 1|X)$, including observations of t and t' and ignoring that we would be pooling together the composition of the treated group in two different points in time. This procedure ignores the fact that the composition of the treated group after the treatment may have changed. Another alternative is to try to balance the distribution of X among the four cells, which means that we must take into account that X is also distributed between t and t'. Blundell et al. (2001) suggest that besides the usual propensity score $P(I = 1|X)$, the probability of being observed in period t (after the treatment), $P(t = 1|X)$, should also be used in the matching process[50].

We apply two matching procedures to our data. Following Blundell et al. (2001), first we use Euclidean Distance to combine the two propensity scores in the search for the closest matching for the NN matching case. Second we estimate a kernel estimator. We multiply the two kernel weights that we get from the two propensity scores for each observation on the treated group, and divide its product by the

[49]Notice that this assumption differs from (5.4) only by the conditioning on the propensity score.

[50]Smith and Todd(2000) put forward a procedure to estimate the common support region. Basically they argue that the intersection of the two densities with some trimming for very low probabilities should be selected. However, the widely used user-written programs in Stata pscore.ado and attk.ado define the common support in a simpler way. Nonetheless they differ between them in the way that the distribution is trimmed. While psmatch.ado excludes those observations in the treated group whose probability (or linear index) is not within the common support, attk.ado excludes those observations in the control group whose probability (or linear index) is not within the common support. In this paper we exclude both treated and control observations that are not within the common support. We believe that this procedure is much more in line with Smith and Todd (2000) idea than the former ones.

product of their constant bandwidth given by $1.06\sigma_x N^{(-0.2)}$.

Another important point to notice is that the matched control before and after and the treated before are found separately, one each time, for each observation on the treated after (group). The NN estimator is calculated by subtracting the differences between the four matched group mean. In the case of the Kernel estimator the weighted mean for each group is calculated for each group and then the differences are taken. The standard errors of the estimators are calculated via 150 bootstrap repetitions.

5.5 Data

The microdata used in this chapter comes from the Brazilian Monthly Employment Survey (PME) published by IBGE (*Instituto Brasileiro de Geografia e Estatistica*). The PME collects monthly data in a similar way of the American CPS. Each household stays in the sample for 4 months, then "rests" for 8 months, and finally returns for extra 4 months. As the unit of analysis is the household, nothing guarantees that the same family(ies) will be living in that address for the whole survey period. Thus, it is fundamental to build an individual identifier that also incorporates individual characteristics in order to follow the individuals over time. Besides the household identifier, we use the serial number of the individual, the week and the month of the interview - those variables are important because they allow us to follow correctly the panel -, the day, the month and the year of birth of the individual as well as gender. Assessing the identifier built in this fashion, we verified that there is no individual who appears more than 8 times in any pair of two years, or more than 4 times within a year. This means that the identifier really avoids the problem of mixing up similar individuals. However, nothing guarantees that the individual will be followed consecutively during the period in which he/she can be interviewed. The individual can drop out definitely, or drop out temporarily and return to the sample later on.

The attrition rate is quite high, therefore, the number of observations reduces considerably if we try to follow an individual for a year - which implies at least 5 observations - or for the 8 possible appearances in the sample. Due to this reason and to the extremely high rate of turnover that characterizes the Brazilian labour market, we will analyse only monthly transitions. The advantage of the PME in this case is that it asks an individual about his/her monthly wage in every interview and not only in the first interview and in the outgoing one as is common in this sort of survey.

A worrying limitation of the PME is that the individual is asked about his/her monthly wage in the last month, and not in the month of the interview. Moreover, if the individual is unemployed or inactive in the month of the interview, he/she is not asked about the wage in the previous month. Thus, even if he/she had a job in $t-1$, we do not have information about his/her wage. For this reason, we do not know the wage of an individual who became unemployed in a month of a minimum wage increase, so we assume that his/her wage in the month previous to the minimum wage increase was equal to his/her wage in the month previous to that. This means that for an unemployed (or inactive) individual in t (month with minimum wage increase), the wage in $t-1$ is assumed, in our analysis, to be equal to the wage in $t-2$. For the individuals who are employed in t we do observe their wages in $t-1$. Another limitation is that we cannot guarantee that the last month wage refers to the wage of the current job, since the individual may have switched job without changing his/her employment status, or even, his/her occupation or sector of activity. As there is no way to control for that, we will assume that the individuals are in the same job for two consecutive months.

Our outcome variables are non-employment[51], unemployment, out of the labour force, non-registered (informal) and self-employed status in t, after the real or the pseudo increase in the minimum wage. For the period 1995 to 1999 the minimum wage was increased every May. So we will compare transitions between April and May ($t = 1$) with transitions for the pair of months immediately before: March to April and for a pair of months 5 months ahead of the last increase, the pair September/October. The treated group is made up of workers who earn between the current minimum wage and the new one, and the comparison groups are made up of workers who earn between the new minimum wage and twice the current one. For the September/October pair the pseudo-increase is equal to the increase that would be observed in May of the following year. Treated individual are the ones who earn between the current minimum wage and the future one, and control group individuals are the ones who earn between the future minimum wage and twice the current one[52].

[51]Non-employment state consists in the the sum of unemployment and out of the labour force states.

[52]See Figure 5.5 for a visual inspection of treated and control groups.

5.6 Results

Before discussing the results of the differences-in-differences estimates of the impact of the minimum wage on the employment transitions, it is worth observing the pattern of monthly employment transitions during the two periods under analysis. We confine our analysis of the empirical transition probabilities to workers between 14 and 65 years. The columns in Tables 5.1 and 5.2 show the employment status in t and the rows the employment status in $t-1$ for the period 1982 – 1986 and 1995 – 1999, respectively. The results are averages over a year. As expected the less mobile groups are formal workers, public servants[53] and inactive individuals. Informal workers are more mobile than self-employed and employers, but less mobile than non-remunerated workers and unemployed individuals. The main difference between the two periods is that both formal and informal workers became more mobile. As for the formal sector, transitions to the informal sector seems to be the main culprit for the fall in the monthly degree of inertia[54]. This change is in line with the increase in the proportion of informal sector workers observed after 1990's and documented in chapter 2 of this thesis. In analyzing the results of the next section it is important to keep in mind that informal workers have a higher mobility than their formal counterparts.

5.6.1 Minimum Wage with Indexation: The period 1982 - 1985

The period 1982-1985 was characterized by a 6-month periodical increase in the nominal value of the minimum wage and by the full indexation of its value to the inflation rate. Table 5.3 shows the nominal values of the minimum wage for all episodical increases under analysis. Until May, 1984 there were two different minimum wages for the metropolitan areas covered by PME - the lowest value was applied to the two northeastern metropolitan areas (Recife and Salvador). For this period we only analyse the differences in employment transitions between the pair of months just before the increase in the minimum wage - March to April and September to October - and the pair of months when the increase took place - April to

[53]The high proportion of public servants coming from formal and informal sector in 1982 seems to be noise in the data, since it is not compatible with the behaviour of this flow in any of the other years analysed. One possible explanations for this difference is that public servants are defined based on a filter of occupations and not on self-declaration of the interviewees (See chapter 2 subsection 2.2.1 for an explanation about the filter). Thus it is possible that this code does not work very well in that specific year.

[54]Notice that transitions from formal status to unemployment are quite stable over the years.

May and October to November. The comparison group is defined as workers who earned between the new minimum wage and two times the old minimum wage. Due to the 100% increase observed in May 1985 we choose a higher upper bound for the

1982	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**90.65**	2.75	2.75	0.79	0.16	0.03	1.32	1.55
informal	10.97	**62.08**	7.86	6.79	0.6	0.7	3.68	7.32
public serv.	12.91	10.72	**74.24**	0.59	0.21	0.06	0.22	1.06
self-employed	2.32	7.29	0.33	**74.4**	3.54	0.52	1.81	9.79
employer	2.19	2.75	0.41	15.11	**76.76**	0.65	0.48	1.64
not remunerated	1.65	9.22	0.4	6.41	1.73	**58.91**	1.88	19.81
unemployed	10.64	15.28	0.55	6.39	0.25	0.47	**39.05**	27.37
inactive	1.01	2.54	0.15	3.23	0.11	0.54	3.12	**89.31**
Total	31.92	10.22	5.02	10.77	2.51	0.9	3.65	35.02
1983	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**90.92**	3.01	1.13	1	0.23	0.04	1.7	1.97
informal	12.12	**62.71**	0.99	8.5	0.67	0.63	4.93	9.45
public serv.	5.01	1.25	**90.89**	0.62	0.22	0.07	0.39	1.55
self-employed	2.56	7.66	0.41	**71.58**	3.58	0.5	2.68	11.02
employer	2.64	2.61	0.56	15.91	**75.2**	0.58	0.6	1.9
not remunerated	1.36	8.23	0.74	6	1.62	**58.76**	1.98	21.31
unemployed	9.46	14.05	0.56	8.14	0.31	0.44	**40.73**	26.31
inactive	1.17	2.55	0.25	3.13	0.13	0.53	3.23	**89.02**
Total	29.53	9.16	6.57	10.49	2.43	0.87	4.25	36.69
1984	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**91.69**	3.15	1.1	0.88	0.2	0.04	1.38	1.56
informal	11.92	**63.68**	0.85	8.71	0.68	0.69	4.83	8.64
public serv.	5.07	1.3	**91.3**	0.58	0.21	0.06	0.3	1.18
self-employed	2.43	8.02	0.39	**70.8**	3.38	0.54	3.15	11.29
employer	2.64	2.89	0.54	15.89	**74.61**	0.63	0.67	2.12
not remunerated	1.69	9.03	0.47	6.59	2.11	**56.34**	2.3	21.46
unemployed	7.27	13.6	0.55	8.93	0.28	0.51	**41.83**	27.03
inactive	1	2.7	0.24	3.56	0.13	0.56	3.88	**87.95**
Total	28.88	9.88	6.46	11.1	2.37	0.89	4.73	35.68
1985	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**90.72**	3.38	1.21	1.16	0.26	0.04	1.3	1.95
informal	13.88	**60.96**	0.97	9.32	0.7	0.61	3.71	9.85
public serv.	5.6	1.37	**90.19**	0.71	0.22	0.08	0.32	1.52
self-employed	3.44	8.58	0.5	**68.82**	3.55	0.58	2.23	12.31
employer	3.08	2.68	0.83	16.86	**73.1**	0.69	0.46	2.3
not remunerated	1.98	9.04	0.79	8.04	2.21	**53.48**	1.82	22.63
unemployed	11.28	14.82	0.87	8.85	0.28	0.39	**35.01**	28.5
inactive	1.48	2.99	0.32	3.72	0.16	0.51	2.96	**87.87**
Total	30.22	9.7	6.57	10.87	2.35	0.78	3.41	36.1
1986	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**90.28**	3.68	1.33	1.23	0.27	0.04	1.11	2.06
informal	15.47	**60.37**	1.01	9.55	0.66	0.48	2.53	9.92
public serv.	5.94	1.48	**89.77**	0.63	0.21	0.09	0.26	1.63
self-employed	3.57	8.85	0.45	**69.27**	3.83	0.5	1.28	12.24
employer	3.23	2.99	0.67	17.49	**72.47**	0.65	0.21	2.3
not remunerated	2.25	9.33	1.11	8.65	2.38	**49.98**	1.03	25.27
unemployed	13.17	14.79	0.87	7.8	0.21	0.36	**32.07**	30.73
inactive	1.57	2.98	0.34	3.63	0.16	0.44	2.16	**88.72**
Total	31.62	9.52	6.77	10.7	2.4	0.62	2.35	36.01

Notes: 1) Initial state correspond to the rows and final state to the columns
2) Transitions from December to January are missing in the calculations

Table 5.1: Monthly Transitions Between Employment Status: 1982-1986

1995	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**87.46**	5.1	1.34	1.86	0.32	0.03	1.25	2.64
informal	16.22	**57.62**	1.31	10.65	0.65	0.45	3.08	10.02
public serv.	5.87	2.11	**88.27**	0.91	0.27	0.04	0.25	2.27
self-employed	3.48	8.23	0.43	**71.98**	3.19	0.44	1.91	10.34
employer	3.19	2.62	0.55	17.57	**72.54**	0.8	0.3	2.43
not remunerated	1.53	8.79	0.58	10.81	2.94	**49.3**	0.99	25.06
unemployed	8.07	13.14	0.44	11.16	0.25	0.31	**36.99**	29.64
inactive	1.62	3.03	0.38	4.06	0.16	0.44	2.37	**87.93**
Total	26.96	10.03	6.22	13.84	2.54	0.61	2.94	36.86
1996	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**86.97**	5.16	1.36	1.95	0.33	0.04	1.34	2.86
informal	15.44	**57.85**	1.47	10.89	0.73	0.47	3.25	9.89
public serv.	5.89	2.35	**87.53**	1	0.27	0.06	0.31	2.6
self-employed	3.32	8.05	0.44	**71.77**	3.23	0.47	2.06	10.65
employer	3.22	3.27	0.55	18.41	**70.61**	0.82	0.48	2.64
not remunerated	1.92	8.58	0.73	10.55	4.17	**47.63**	1.36	25.06
unemployed	8.1	12.59	0.59	10.93	0.3	0.27	**36.36**	30.86
inactive	1.65	3.18	0.41	4.11	0.16	0.46	2.74	**87.31**
Total	25.87	10.3	6.31	14.27	2.59	0.63	3.32	36.72
1997	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**87.47**	5.23	1.32	1.79	0.28	0.02	1.19	2.69
informal	15	**58.21**	1.38	11.06	0.74	0.43	3.29	9.89
public serv.	5.37	2.26	**88.64**	0.81	0.19	0.06	0.3	2.38
self-employed	3.14	8.14	0.39	**71.88**	3.27	0.46	2.31	10.41
employer	2.65	3.16	0.54	18.07	**71.93**	0.89	0.45	2.3
not remunerated	1.78	9.76	0.39	11.68	3.73	**46.67**	1.6	24.39
unemployed	6.62	11.99	0.5	10.54	0.27	0.27	**37.6**	32.21
inactive	1.45	2.96	0.39	3.88	0.16	0.4	3	**87.76**
Total	25.37	10.19	6.23	14.07	2.56	0.58	3.51	37.5
1998	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**86.88**	5.06	1.4	1.75	0.3	0.03	1.55	3.02
informal	14.29	**57.88**	1.44	10.57	0.8	0.5	4.12	10.39
public serv.	5.6	2.44	**87.56**	1.03	0.29	0.06	0.37	2.64
self-employed	3.06	8.22	0.48	**70.69**	3.37	0.46	3.04	10.69
employer	3.04	3.22	0.63	18.29	**70.53**	0.78	0.64	2.87
not remunerated	1.55	10.14	0.77	11.91	3.26	**45.41**	1.93	25.03
unemployed	6.32	11.53	0.56	10.09	0.2	0.27	**39.91**	31.11
inactive	1.53	3.06	0.41	3.84	0.16	0.39	3.92	**86.69**
Total	24.59	10.27	6	13.6	2.49	0.57	4.69	37.78
1999	formal	informal	public serv.	self-employed	employer	non-payed	unemp.	inactive
formal	**87.35**	5.26	1.33	1.68	0.29	0.03	1.23	2.83
informal	13.45	**59.39**	1.56	10.5	0.79	0.49	3.93	9.89
public serv.	5.31	2.45	**88.09**	0.87	0.22	0.06	0.44	2.57
self-employed	2.92	8.1	0.47	**71.13**	3.14	0.45	3.13	10.65
employer	2.93	3.42	0.6	17.8	**70.83**	0.88	0.63	2.91
not remunerated	1.41	10.78	0.97	10.18	4.12	**45.47**	1.98	25.08
unemployed	5.04	10.38	0.56	10.18	0.26	0.26	**41.34**	31.99
inactive	1.38	2.84	0.37	3.66	0.17	0.37	3.77	**87.43**
Total	23.62	10.35	6.04	13.5	2.45	0.55	4.64	38.85

Notes: 1) Initial state correspond to the rows and final state to the columns
2) Transitions from December to January are missing in the calculations

Table 5.2: Monthly Transitions Between Employment Status: 1995-1999

comparison group: workers who earned three times the old minimum wage[55].

[55]Notice that an 100% increase in nominal terms means that we would end up without observations in the control group, since we only include in the control groups workers who earned up to two times the current minimum wage.

	Minimum Wage	Variation	Minimum Wage - NE	Variation
Apr-82	11928		10200	
May-82	16608	39%	14400	41%
Nov-82	23568	42%	20736	44%
May-83	34776	48%	30600	48%
Nov-83	57120	64%	50256	64%
May-84	97176	70%	97176	93%
Nov-84	166560	71%	166560	71%
May-85	333120	100%	333120	100%
Nov-85	600000	80%	600000	80%
Apr-95	70		70	
May-95	100	43%	100	43%
May-96	112	12%	112	12%
May-97	120	7%	120	7%
May-98	130	8%	130	8%
May-99	136	5%	136	5%

Notes: 1) Until May 1984 the metropolitan areas in the Northeast (NE) had a lower minimum wage; 2) Currencies: *Cruzeiros* until November 1985 and *Reais* from April 1995 onwards.

Table 5.3: Nominal Minimum Wage and Nominal Increases

Probit Analysis

All probit regressions were run with the following controls: gender, age, age squared, six education groups[56], region, and industry. The coefficients of interest is the interaction between the month of the minimum wage increase and the treated group ($t{*}m$) that gives us the sign of the difference in difference, i.e, whether or not workers affected by the minimum wage increases had a different (worse) transition pattern in the labour market in the month of the increase when compared to similar comparison groups. A positive and significant coefficient indicates that treated observations were affected by the minimum wage in the expected direction, i.e, the minimum wage hike caused disemployment or transitions to the informal sector or to self-employment, according to the specification in use.

Table 5.4 brings the results for the pooled sample of formal and informal workers and for each group separately of the marginal effect of the interaction term of the dummy for the treated group and the dummy for the month after the minimum wage increase $t{*}m$[57]. For the nonemployment equation there is no clear pattern for the sign of the marginal effect of interaction term: two out of eight coefficients display an unexpected negative sign, but only the one for May 1984 is statistically significant at 5%, which means that after the minimum wage increase there was a reduction in transitions to nonemployment for treated workers when compared to the comparison group in that episode. The other 6 positive estimates are not statistically

[56]The same education groups used in the previous chapters.

[57]Results not reported here show that the marginal effects of the dummy for the treated group tend to be positive and significant (but not always), this result means that workers between the old and the new minimum wage are more likely to move to nonemployment than the control group even without the presence of the treatment, i.e, even without the minimum wage hike.

	Dep. Variable	Pooled Sample			Formal			Informal		
			t*m	z		t*m	z		t*m	z
1982	Transitions to:									
	Non-employment	May	-0.003	[0.43]	May	-0.002	[0.19]	May	-0.016	[0.74]
	Unemployment	(N=16369)	-0.001	[0.12]	(N=12720)	-0.001	[0.21]	(N=3647)	-0.003	[0.22]
	Inactivity		-0.002	[0.35]		0	[0.04]		-0.011	[0.69]
	Non-employment	Nov	0.000	[0.01]	Nov	0.004	[0.54]	Nov	-0.014	[0.67]
	Unemployment	(N=18815)	-0.006	[1.37]	(N=15063)	-0.003	[0.56]	(N=3752)	-0.019	[1.66]
	Inactivity		0.006	[1.18]		0.006	[1.20]		0.008	[0.48]
1983	Transitions to:									
	Non-employment	May	0.01	[1.26]	May	0.002	[0.28]	May	0.037	[1.55]
	Unemployment	(N=17974)	0.006	[1.16]	(N=14424)	0.002	[0.39]	(N=3550)	0.021	[1.52]
	Inactivity		0.004	[0.74]		0.001	[0.11]		0.018	[1.03]
	Non-employment	Nov	0.016	[1.55]	Nov	0.007	[0.76]	Nov	0.044	[1.42]
	Unemployment	(N=17974)	0.001	[0.10]	(N=13965)	0.001	[0.11]	(N=3581)	0.003	[0.20]
	Inactivity		0.017	[2.08]*		0.007	[0.98]		0.047	[1.81]
1984	Transitions to:									
	Non-employment	May	-0.023	[2.08]*	May	-0.003	[0.23]	May	-0.084	[2.70]**
	Unemployment	(N=16672)	-0.008	[1.16]	(N=14389)	0.004	[0.60]	(N=3395)	-0.05	[2.76]**
	Inactivity		-0.013	[1.50]		-0.008	[0.94]		-0.022	[0.96]
	Non-employment	Nov	0.015	[1.23]	Nov	0.011	[1.03]	Nov	0.021	[0.48]
	Unemployment	(N=17202)	0.005	[0.72]	(N=13630)	0.007	[0.94]	(N=3572)	-0.008	[0.36]
	Inactivity		0.008	[0.84]		0.003	[0.36]		0.051	[1.30]
1985	Transitions to:									
	Non-employment	May	0.001	[0.12]	May	-0.008	[1.22]	May	0.068	[2.30]*
	Unemployment	(N=24828)	-0.002	[0.35]	(N=20951)	-0.005	[1.10]	(N=3877)	0.021	[1.17]
	Inactivity		0.002	[0.48]		-0.003	[0.67]		0.043	[1.92]
	Non-employment	Nov	0.005	[0.43]	Nov	0.009	[0.82]	Nov	-0.013	[0.43]
	Unemployment	(N=17557)	-0.001	[0.13]	(N=13949)	0.003	[0.49]	(N=3608)	-0.012	[0.97]
	Inactivity		0.005	[0.54]		0.006	[0.68]		-0.002	[0.07]

*Significant at 5%, **Significant at 1%.
Robust z statistics in brackets

Table 5.4: Changes in Transition to Non-employment: Probit (1982-1985)

significant. All in all, it seems that minimum wage increases had no impact on the probability of moving from employment to nonemployment in the period 1982 - 1985, when using this specific comparison group. Not surprisingly, the estimates for both the probability of transition to unemployment and the probability of transition to inactivity (out of the labour force) do not show any clear effect either. Only the increase of the minimum wage in November 1983 seems to have had a positive and significant impact on transitions to inactivity (1.7%).

As for the sample of formal sector workers, the results are quite similar to the previous one for the whole sample. There is no evidence that minimum wage hikes increase the probability that treated workers move from employment to nonemployment relative to their control counterparts.

As for the informal sector workers[58], we find significant marginal effects for the interaction term in May 1984 and May 1985. However, whereas for 1984 the coefficient is negative, indicating that the minimum wage increase made informal treated workers more employable than their control counterparts, in 1985 it is positive in-

[58]Similarly to the pooled sample and the sample for formal workers, results not reported here show that in general treated workers are more likely to move to non-employment than their counterparts in the control group. However, this result is not as strong as for the former samples.

		Formal		
	Dep. Variable		t^*m	se
1982	Transitions to:			
	Informal Sector	May	0.008	[1.02]
	Self-Employment	(N=12720)	-0.002	[2.91]**
	Informal Sector	Nov	-0.007	[1.23]
	Self-Employment	(N=15063)	0.000	[0.20]
1983	Transitions to:			
	Informal Sector	May	-0.004	[0.68]
	Self-Employment	(N=14424)	-0.001	[0.68]
	Informal Sector	Nov	0.006	[0.69]
	Self-Employment	(N=13965)	0.002	[0.74]
1984	Transitions to:			
	Informal Sector	May	0.014	[1.27]
	Self-Employment	(N=14389)	-0.004	[1.25]
	Informal Sector	Nov	-0.021	[2.02]*
	Self-Employment	(N=13630)	-0.001	[0.36]
1985	Transitions to:			
	Informal Sector	May	0.001	[0.07]
	Self-Employment	(N=20951)	0.004	[1.16]
	Informal Sector	Nov	-0.026	[2.81]**
	Self-Employment	(N=13949)	0.000	[0.13]

*Significant at 5%, **Significant at 1%.

Table 5.5: Changes in Transition to Informality: Probit (1982-1985)

dicating that there was an increase in transition to non-employment for treated individuals[59]. But again there is no pattern in the sign of the interactions during this period.

As for the transitions to informality the only significant interaction terms were found in November 1984 and November 1985 and with a negative sign, implying that the minimum wage hike led to a reduction in transitions to the informal sector (non-registered state). As for transition to self-employment, only the interaction term for May 1982 is negative and statistically significant.

All in all, the results for this sample period indicate a negligible effect of minimum wage hikes on employment transitions. This conclusion is reinforced by Tables C.1, C.2, and C.3 in the Appendix that show the probit estimates for minor changes in the way we define the control groups. We test the sensitivity of the above results to the inclusion of workers whose wage may be potentially indexed to the minimum wage, for that reason we exclude from the sample workers who earn exactly 1.5 and 2 minimum wages[60]. The results for this alternative definition of treated and control groups are reported in Tables C.1 and C.3. We also rerun the probit estimates defining treated and control groups based on the wage in $t-2$, instead of $t-1$ for employed individuals in t[61]. The results are reported in Ta-

[59]It is interesting to note that whereas the positive effect on nonemployment in 1985 seems to be determined by an increase in transitions to inactivity, the negative effect on nonemployment of 1984 was determined by a reduction in transitions to unemployment

[60]For May 1985 we also exclude workers who earn exactly 3 minimum wages.

[61]Actually, we are generalizing the assumption that we made for unemployed individuals in t for

bles C.2 and C.3. The results are basically the same of the previous specifications. No effect at all for the formal sector sample, and some effect without a clear pattern for informal sector workers.

Matching: Kernel-based results

The specification used to estimate the propensity score for both the treated group $P(I = 1|X)$ and for the time after treatment $P(t = 1|X)$ - to correct for compositional change[62] - was the same that we used for the probit estimates.

The estimates of interest here are the differences between treated and control groups in the probability of moving to the nonemployment, unemployment, inactivity, informality and self-employment in the months after minimum wage increases when compared to months when the increase did not take place.

As in the case of the probit estimates, the propensity score matching estimates for the pooled sample do not exhibit a common pattern on the effect of nominal minimum wage hikes on transitions to nonemployment. Table 5.6 shows that the sign of the differences in probabilities varies considerably. For November 1982 and May 1984 it seems to have occurred a decrease in the probability of nonemployment for treated workers, whereas for the other episodes over this period there has been an increase. But this result is only statistically significant for May 1984. This result is similar to the one found by the probit estimates of the marginal effect of the difference-in-differences approach (Table 5.4). As for transitions to unemployment only the estimates for May 1983 are positive and significant and for transitions to inactivity there is a statistically significant decrease in May 1984.

As for the sample of formal workers, most of the estimates of differences in the probability of transition to nonemployment are positive, but there are negative signs for November 1982 and May 1984. Nevertheless none of the estimates are statistically significant like in the probit specifications. The same lack of pattern occurs again for the decomposition of this transitions into unemployment and inactivity.

As for the sample of informal workers, most of the estimates of the difference in the transitions to nonemployment are positive, but there are some negative estimates: May 1982, May 1984 and November 1985. However, only the estimate for May 1984[63] is statistically significant, which means that the minimum wage hike of

whom we do not observe the wage in $t - 1$.

[62] We do not report the results of the probit estimations used to calculate the propensity score indexes used in the matching process. But it is worth mentioning that in general the covariates were not significant in the equation for $P(t = 1|X)$, so it seems the compositional effects are not a major issue in this particular application.

[63] Again this result is in line with the probit estimates.

	Difference in Differences in Probability					
	Nonemployment		Unemployment		Inactivity	
Pooled Sample	estimates	se	estimates	se	estimates	se
May-82	0.005	[0.008]	0.001	[0.005]	0.004	[0.006]
Nov-82	0.000	[0.006]	-0.007	[0.004]	0.007	[0.005]
May-83	0.011	[0.007]	0.011	[0.005]**	0.000	[0.005]
Nov-83	0.013	[0.008]	0.006	[0.006]	0.007	[0.006]
May-84	-0.024	[0.009]**	-0.008	[0.007]	-0.015	[0.006]**
Nov-84	0.009	[0.009]**	0.006	[0.007]	0.003	[0.005]
May-85	0.012	[0.007]	0.002	[0.004]	0.010	[0.006]
Nov-85	0.007	[0.010]	0.006	[0.006]	0.001	[0.007]
Formal Sector						
May-82	0.002	[0.008]	-0.001	[0.006]	0.003	[0.006]
Nov-82	-0.004	[0.081]	-0.007	[0.005]	0.003	[0.005]
May-83	0.009	[0.079]	0.010	[0.006]	-0.001	[0.006]
Nov-83	0.010	[0.009]	0.004	[0.006]	0.005	[0.007]
May-84	-0.009	[0.011]	0.001	[0.007]	-0.010	[0.008]
Nov-84	0.003	[0.008]	0.004	[0.007]	-0.002	[0.006]
May-85	0.006	[0.007]	0.002	[0.004]	0.004	[0.005]
Nov-85	0.005	[0.008]	0.002	[0.006]	0.002	[0.082]
Informal Sector						
May-82	-0.017	[0.021]	-0.013	[0.015]	-0.004	[0.017]
Nov-82	0.020	[0.024]	-0.009	[0.013]	0.029	[0.020]
May-83	0.041	[0.025]	0.032	[0.017]	0.009	[0.019]
Nov-83	0.043	[0.027]	0.004	[0.207]	0.039	[0.021]
May-84	-0.090	[0.029]**	-0.057	[0.021]**	-0.033	[0.023]
Nov-84	0.008	[0.029]	-0.018	[0.017]	0.026	[0.020]
May-85	0.024	[0.102]	-0.001	[0.012]	0.025	[0.022]
Nov-85	-0.023	[0.022]	0.001	[0.013]	-0.024	[0.017]

*Significant at 5%, **Significant at 1%.

Table 5.6: Changes in Transition to Non-Employment: Kernel Propensity Score Matching (1982-1985)

	Difference in Differences in Probability			
	Informality		Self-employment	
Formal Sector	estimates	se	estimates	se
May-82	0.002	[0.008]	-0.005	[0.002]*
Nov-82	0.003	[0.006]	-0.001	[0.002]
May-83	0.001	[0.007]	0.001	[0.003]
Nov-83	0.002	[0.008]	0.001	[0.002]
May-84	0.011	[0.011]	-0.004	[0.011]
Nov-84	-0.016	[0.008]*	-0.001	[0.003]
May-85	-0.006	[0.006]	-0.003	[0.003]
Nov-85	-0.027	[0.009]**	0.007	[0.004]

*Significant at 5%, **Significant at 1%.

Table 5.7: Changes in Transition to Informality: Kernel Propensity Score Matching (1982-1985)

May 1984 led to a lower transition to unemployment to treated informal workers when compared to the control informal workers.

As with the probit estimates, the kernel estimates of changes in the probability of transition to the informal sector reported in Table 5.7 only show a statistically significant effect of the minimum wage hike for November 1984 and November 1985 with reductions in this outflow. There is also an increase in transitions to self-employment in November 1985.

Overall, the results for this period using the kernel estimates do not diverge much from the result of the probit analysis and point to the absence of negative effect of the minimum wage hikes on employment.

Matching: Nearest Neighbour-based Results

The nearest neighbour propensity score matching results are also based on the two propensity scores $P(I = 1|X)$ and $P(t = 1|X)$ with the same specification as the one used for the probit. Unlike the kernel-based matching, the nearest neighbour uses only one member of the control group to generate the counterfactual for the treated group.

The results of the NN estimates for the pooled sample in Table 5.8 are positive and significant for November 1983 and 1984 and May 1985. Only in May 1984 do we find a negative, but statistically insignificant effect. However the results are very different from the ones the kernel matching and of the probit estimates. Only the positive effect on the probability of moving to inactivity in November 1983 was also found in the probit analysis. The other significant results refer to an increase in transitions to inactivity also in November 1983 and 1984, and May 1985.

The results for the formal workers sample also show a pattern of positive transitions to nonemployment (again with the exception of May 1984), but only the estimate for November 1983 and May 1985 are statistically significant.

The results for the informal workers, in general, show a positive effect on the difference in the probabilities for the treated group. This result does not hold for May 1984 and November 1985. The results are statistically significant for May 1983 and May 1984. However, whereas in May 1983 the minimum wage hike led to an increase in transitions to nonemployment, the increase in May 1984 had the opposite effect. Notice that the result for May 1984 is in line with the probit estimates.

As for transitions to informality and self-employment the results in table 5.9 suggest that for both destinations the difference in probability is in general positive. However, most of the estimates are not statistically significant, the exception being November 1983 and November 1985 for the self-employment equation indicating an increase in transitions to self-employment[64].

[64]This latter result is in line with the kernel estimates.

	Difference in Differences in Probability					
	Nonemployment		Unemployment		Inactivity	
Pooled Sample	estimates	se	estimates	se	estimates	se
May-82	0.018	[0.010]	0.009	[0.007]	0.009	[0.008]
Nov-82	0.005	[0.009]	-0.008	[0.005]	0.013	[0.007]
May-83	0.015	[0.010]	0.015	[0.007]	0.000	[0.008]
Nov-83	0.025	[0.011]*	0.005	[0.008]	0.020	[0.007]**
May-84	-0.015	[0.010]	-0.006	[0.007]	-0.009	[0.081]
Nov-84	0.023	[0.011]*	0.009	[0.009]	0.014	[0.007]*
May-85	0.019	[0.008]*	0.002	[0.005]	0.017	[0.006]**
Nov-85	0.011	[0.011]	0.009	[0.007]	0.002	[0.008]
Formal Sector						
May-82	0.007	[0.011]	0.007	[0.007]	0.001	[0.009]
Nov-82	0.003	[0.008]	0.000	[0.006]	0.004	[0.005]
May-83	0.007	[0.010]	0.008	[0.007]	-0.001	[0.008]
Nov-83	0.026	[0.012]*	0.005	[0.008]	0.020	[0.008]*
May-84	-0.004	[0.011]	0.007	[0.006]	-0.011	[0.010]
Nov-84	0.017	[0.011]	0.012	[0.009]	0.005	[0.006]
May-85	0.021	[0.010]*	0.007	[0.006]	0.014	[0.008]
Nov-85	0.023	[0.082]	0.010	[0.007]	0.014	[0.008]
Informal Sector						
May-82	0.023	[0.024]	-0.005	[0.078]	0.027	[0.020]
Nov-82	0.035	[0.030]	-0.006	[0.081]	0.041	[0.025]
May-83	0.066	[0.029]*	0.039	[0.018]*	0.027	[0.022]
Nov-83	0.048	[0.030]	-0.022	[0.022]	0.069	[0.022]*
May-84	-0.068	[0.032]*	-0.048	[0.023]*	-0.020	[0.026]
Nov-84	0.024	[0.023]	-0.012	[0.015]	0.037	[0.016]*
May-85	0.031	[0.026]	-0.015	[0.015]	0.047	[0.023]
Nov-85	-0.018	[0.024]	-0.004	[0.015]	-0.014	[0.017]

*Significant at 5%, **Significant at 1%.

Table 5.8: Changes in Transition to Non-employment: Nearest Neighbour Score Matching (1982 - 1985)

	Difference in Differences in Probability			
	Informality		Self-employment	
Pooled Sample	estimates	se	estimates	se
May-82	0.005	[0.010]	0.000	[0.002]
Nov-82	0.008	[0.009]	0.005	[0.003]
May-83	0.000	[0.008]	-0.005	[0.003]
Nov-83	0.002	[0.009]	0.007	[0.003]**
May-84	0.014	[0.011]	-0.005	[0.010]
Nov-84	-0.002	[0.008]	0.001	[0.003]
May-85	-0.011	[0.008]	-0.001	[0.003]
Nov-85	-0.020	[0.011]	0.012	[0.005]*

*Significant at 5%, **Significant at 1%.

Table 5.9: Changes in Transition to Informality: Nearest Neighbour Score Matching (1982-1985)

5.6.2 Minimum wage without indexation: the period 1995-1999

The period 1995-1999 witnessed five minimum wage increases. Unlike the period 1982-1985, the minimum wage increased during this period in real terms. The impact of the increase in real terms, nonetheless, varied a lot from year to year, but the biggest real increase was in May 1995 as can be seen in Figure 5.1.

Besides using the pair of months just before the increase in the minimum wage

- March/April - as contrast for the pair of months of the increase - April/May, we simulate an increase in the minimum wage between September and October and contrast the effect of this pseudo-increase with the real increase of April/May. In the pseudo-increase for the pair September/October the treated group is defined as workers earning between the current minimum wage and the minimum wage of May of the following year, and the control group is defined as workers earning between the future minimum wage and two times the current one[65].

Probit Analysis

Table 5.10 shows that the marginal effect of the interaction t^*m alternates positive and negative values when the pair of months immediately before the minimum wage increase is used March/April (*May*), but it is always positive when the pair September/October (*Oct*) is used as contrast. However for most specification the estimates are not significant. Only for 1996 and 1998 the interaction is statistically significant and only when the pair September/October is used as contrast[66]

Decomposing the impact on non-employment into its two components, unemployment and inactivity, we find even more volatility in the results with the sign of the interaction terms varying a lot. The significance of the positive interaction term for 1996[67] and 1998 discussed above is driven by the transitions to unemployment, which are positive and significant when the pair September/October is used as contrast. As for the effect on the probability of moving to inactivity the results of the interaction term were never significant.

As for the formal sector[68], the interaction term in the nonemployment equation, displays in general positive coefficients as expected, but they are never statistically significant. Interestingly, the large nominal (and also real) increase in the minimum wage observed in 1995 led to a negative (but not significant) sign for the interaction term when the pair of months just before the increase is used as contrast. Looking

[65]Notice that the results for October are not a different minimum wage increase as were the results for November for the period 1982-1985. The results reported on the rows *Oct* refer to the comparison between transitions from April to May and transitions from September to October.

[66]As it was the case for the period 1982-1985, estimates not shown here reveal that the treated group is in general more likely to move to nonemployment than the control group regardless of minimum wage hikes. Transitions to inactivity rather than transitions to unemployment is the major force driving this result.

[67]In this case the estimate is significant only at 10%.

[68]Results not reported here show that the treated group does not necessarily have a higher probability of moving to non-employment than the comparison group. This result is only observed for 1995 regardless of the pair of months used as control, whereas it was observed for most of the years in the period 1982-1985 and for the pooled sample as discussed above. Only transitions from employment to inactivity seems to be more likely for the treated group than in the control, even when the coefficient is not significant, it is always positive.

		Pooled Sample			Formal Sector			Informal Sector		
	Dep. Variable		$t^{*}m$	z		$t^{*}m$	z		$t^{*}m$	z
1995	Transitions to:									
	Non-employment	May	-0.008	[0.73]	May	-0.002	[0.14]	May	-0.017	[0.72]
	Unemployment	(N=11948)	-0.004	[0.86]	(N=7714)	-0.001	[0.10]	(N=4234)	-0.007	[0.86]
	Inactivity		-0.003	[0.34]		0.000	[0.03]		-0.011	[0.49]
	Non-employment	Oct	0.007	[0.60]	Oct	0.004	[0.31]	Oct	0.014	[0.60]
	Unemployment	(N=13206)	0.000	[0.06]	(N=8456)	0.004	[0.61]	(N=4750)	-0.004	[0.54]
	Inactivity		0.007	[0.65]		0.000	[0.01]		0.019	[0.89]
1996	Transitions to:									
	Non-employment	May	0.003	[0.24]	May	0.002	[0.15]	May	0.01	[0.42]
	Unemployment	(N=12530)	0.004	[0.65]	(N=7762)	0.005	[0.87]	(N=4704)	0.007	[0.59]
	Inactivity		-0.002	[0.18]		-0.002	[0.22]		0.003	[0.16]
	Non-employment	Oct	0.027	[2.03]*	Oct	0.001	[0.07]	Oct	0.06	[2.35]*
	Unemployment	(N=12084)	0.012	[1.84]	(N=7447)	0.001	[0.24]	(N=4637)	0.021	[1.78]
	Inactivity		0.015	[1.34]		-0.003	[0.28]		0.039	[1.76]
1997	Transitions to:									
	Non-employment	May	0.017	[1.14]	May	0.014	[0.95]	May	0.016	[0.58]
	Unemployment	(N=10949)	-0.003	[0.43]	(N=6958)	0.015	[1.76]	(N=3991)	-0.014	[1.47]
	Inactivity		0.021	[1.59]		0.005	[0.42]		0.04	[1.51]
	Non-employment	Oct	0.009	[0.63]	Oct	0.005	[0.37]	Oct	0.021	[0.78]
	Unemployment	(N=11498)	-0.002	[0.24]	(N=7146)	0.005	[0.58]	(N=4352)	-0.007	[0.64]
	Inactivity		0.009	[0.81]		0.000	[0.01]		0.03	[1.25]
1998	Transitions to:									
	Non-employment	May	0.018	[1.27]	May	0.001	[0.10]	May	0.034	[1.27]
	Unemployment	(N=11399)	0.013	[1.57]	(N=6938)	0.001	[0.20]	(N=4461)	0.025	[1.59]
	Inactivity		0.005	[0.41]		-0.001	[0.08]		0.009	[0.41]
	Non-employment	Oct	0.035	[2.40]*	Oct	0.004	[0.27]	Oct	0.066	[2.39]*
	Unemployment	(N=11915)	0.02	[2.14]*	(N=7331)	0.009	[1.15]	(N=4524)	0.023	[1.49]
	Inactivity		0.015	[1.31]		-0.004	[0.44]		0.042	[1.77]
1999	Transitions to:									
	Non-employment	May	-0.005	[0.35]	May	0.021	[1.25]	May	-0.029	[1.17]
	Unemployment	(N=11300)	-0.006	[0.90]	(N=7057)	0.001	[0.19]	(N=4243)	-0.014	[1.33]
	Inactivity		0.001	[0.12]		0.021	[1.41]		-0.012	[0.56]
	Non-employment	Oct	0	[0.03]	Oct	0.005	[0.35]	Oct	-0.008	[0.31]
	Unemployment	(N=11515)	0.002	[0.28]	(N=7244)	0.006	[0.95]	(N=4271)	-0.004	[0.29]
	Inactivity		-0.001	[0.12]		0.000	[0.04]		-0.003	[0.16]

*Significant at 5%, **Significant at 1%.
Robust z statistics in brackets

Table 5.10: Changes in Transition to Non-employment: Probit (1995-1999)

at the equations for unemployment and inactivity it is striking that the positive - but not significant - interaction terms observed for the nonemployment equation can be attributed to the effect of minimum wage hikes on unemployment. The coefficients of the interaction term on the inactivity equation display a much more volatile behaviour, changing from positive to negative, from year to year, whereas the non-employment equation display positive coefficients for all years, with the exception of 1995 when the pair March/April is used as contrast.

For the sample of informal workers, the results are much more in line with the ones found for the pooled sample than the ones found for formal workers. The coefficients of the interaction terms where positive for all years with the exception of 1995 (when using March/April as contrast) and 1999 (for both pairs of contrast). However, the most interesting fact is that for the year 1996 and 1998 when the pair September/October is used as contrast the positive coefficient is also statistically significant at 5%. For both years, this result seems to be driven by positive changes in the probability of moving to inactivity rather than moving to unemployment. Hence, according to the probit estimates, informal sector workers were the ones who were hit by minimum wage hikes after 1995.

As for the probability of moving to the informal sector (non-registered workers) Table 5.11 shows that again there is no pattern. The interaction assumes positive and negative values, in general statistically insignificant, from one year to another. Only in 1995 and in 1998, when the pair September/October is used as contrast, we do find positive and significant effect of the minimum wage hike on the probability of moving to the informal sector[69]. The same lack of pattern also occurs with the probability of moving to self-employment. In 1996, the minimum wage hike seems to have increased the probability of becoming self-employed[70], while in 1999 it had the opposite effect[71].

The robustness checks for different comparison groups display in Tables C.1 to C.1 in the appendix show that the results for the pooled sample and for the informal sector are robust[72], but the results for the formal sector are less clear. For instance, the positive effect on transition to the informal sector in 1995 and 1998 are not robust to any of the changes in the control group. Nevertheless, there is no indication that minimum wage hikes caused disemployment effect for formal workers. For the pooled sample and for the informal sector sample disemployment

[69] All negative estimates are non-significant.

[70] The coefficient is significant at 10%.

[71] Both significant effects are only found when the pair September/October is used as contrast.

[72] Actually, the results for the informal sector are much more robust than the ones for the pooled sample, since the latter just hold at 10% of significance.

		Formal Sector		
	Dep. Variable		t^*m	z
1995	Transitions to:			
	Informal Sector	May	-0.007	[0.50]
	Self-Employment	(N=7714)	-0.004	[0.70]
	Informal Sector	Oct	0.032	[2.21]*
	Self-Employment	(N=8456)	-0.007	[1.79]
1996	Transitions to:			
	Informal Sector	(N=7762)	-0.011	[0.73]
	Self-Employment	Oct	-0.002	[0.36]
	Informal Sector	(N=7447)	-0.015	[1.04]
	Self-Employment	(N=13965)	0.021	[1.91]
1997	Transitions to:			
	Informal Sector	May	-0.002	[0.11]
	Self-Employment	(N=6958)	-0.001	[0.14]
	Informal Sector	Oct	-0.023	[1.50]
	Self-Employment	(N=7146)	-0.004	[0.77]
1998	Transitions to:			
	Informal Sector	May	0.021	[1.16]
	Self-Employment	(N=6938)	0.006	[0.83]
	Informal Sector	Oct	0.042	[2.38]*
	Self-Employment	(N=7331)	0.015	[1.68]
1999	Transitions to:			
	Informal Sector	May	-0.009	[0.54]
	Self-Employment	(N=7057)	-0.008	[1.60]
	Informal Sector	Oct	0.003	[0.15]
	Self-Employment	(N=7244)	-0.008	[2.57]*

*Significant at 5%, **Significant at 1%.

Table 5.11: Changes in Transition to Informality: Probit (1995-1999)

effects were observed in 1996 and 1998 when the pair September/October are used as contrast.

Matching: Kernel-based results

According to Table 5.12, the estimates using propensity score matching kernel for both contrast pairs, March/April and September/October indicate a negative and insignificant impact of minimum wage on transitions to non-employment for the pooled sample in 1995. For all other years the estimates show a positive effect of the minimum wage hikes on non-employment. However, only for 1996 and 1998 when the pair September/October is used as control this positive impact on non-employment is significant. This result is in line with the probit estimates for the 1998 minimum wage increase. As for the decomposition on transitions to unemployment, the estimates are in general positive, but only in 1998 is the estimate statistically significant and positive (regardless of the pair of months used as contrast[73]). As for transition to inactivity there is a positive and significant effect only in 1997 when March/April is used as contrast.

[73]In the probit analysis this result holds only when the pair September/October is used as contrast.

	Difference in Differences in Probability					
	Nonemployment		Unemployment		Inactivity	
Pooled Sample	estimates	se	estimates	se	estimates	se
May-95	-0.008	[0.012]	-0.003	[0.006]	-0.004	[0.011]
Oct-95	-0.003	[0.012]	0.004	[0.006]	-0.007	[0.011]
May-96	0.010	[0.014]	0.008	[0.006]	0.002	[0.012]
Oct-96	0.028	[0.014]*	0.010	[0.007]	0.018	[0.013]
May-97	0.037	[0.017]*	0.003	[0.007]	0.034	[0.017]*
Oct-97	0.015	[0.016]	0.004	[0.008]	0.010	[0.015]
May-98	0.041	[0.016]**	0.015	[0.008]*	0.026	[0.014]
Oct-98	0.035	[0.015]**	0.019	[0.007]**	0.016	[0.014]
May-99	0.021	[0.016]	-0.007	[0.007]	0.028	[0.016]
Oct-99	0.003	[0.014]	0.005	[0.008]	-0.003	[0.014]
Formal Sector						
May-95	-0.006	[0.011]	0.002	[0.006]	-0.008	[0.108]
Oct-95	-0.005	[0.015]	0.005	[0.007]	-0.011	[0.013]
May-96	0.012	[0.013]	0.006	[0.006]	0.006	[0.013]
Oct-96	0.033	[0.016]*	0.010	[0.008]	0.023	[0.014]
May-97	0.021	[0.019]	-0.001	[0.010]	0.021	[0.017]
Oct-97	0.001	[0.021]	0.002	[0.010]	-0.001	[0.019]
May-98	0.020	[0.018]	0.005	[0.009]	0.0152	[0.016]
Oct-98	0.013	[0.020]	0.014	[0.012]	0.000	[0.016]
May-99	0.015	[0.015]	-0.003	[0.009]	0.018	[0.019]
Oct-99	0.008	[0.019]	0.009	[0.009]	-0.001	[0.016]
Informal Sector						
May-95	-0.002	[0.026]	0.000	[0.014]	-0.002	[0.023]
Oct-95	-0.007	[0.025]	0.001	[0.012]	-0.008	[0.023]
May-96	0.020	[0.029]	0.011	[0.011]	0.009	[0.027]
Oct-96	0.006	[0.026]	0.018	[0.010]	-0.012	[0.023]
May-97	0.059	[0.031]	0.005	[0.013]	0.053	[0.028]
Oct-97	0.029	[0.033]	0.011	[0.018]	0.019	[0.029]
May-98	0.037	[0.023]	0.024	[0.013]	0.013	[0.021]
Oct-98	0.061	[0.029]*	0.017	[0.012]	0.044	[0.026]
May-99	0.013	[0.030]	-0.005	[0.016]	0.018	[0.027]
Oct-99	0.009	[0.029]	-0.004	[0.015]	0.013	[0.026]

*Significant at 5%, **Significant at 1%.

Table 5.12: Changes in Transition to Non-Employment: Kernel Propensity Score Matching (1995-1999)

As for the results for the formal sector, they are quite similar to the ones for the pooled sample in terms of sign: negative effects on transitions to nonemployment for 1995 and positive effects for all other years, regardless of the pair of months used as contrast. However, only for 1996 when the pair September/October is used as contrast is the positive estimate significant. None of the estimates for transitions to unemployment and for transitions to inactivity are significant.

The same pattern of transitions to nonemployment is found in the case of informal workers. Only for 1995 there was a reduction in transitions to nonemployment. For all other years the estimates are positive. However, as was the case for the pooled sample, only for 1996 and when the pair September/October is used as contrast the positive estimate is significant. Moreover, none of the estimates for transitions to unemployment and for transitions to inactivity are significant.

As for transitions to informality, Table 5.13 shows that there is no pattern in the sign of the change. The only statistically significant estimate is a negative effect for 1997 when the pair September/October is used as contrast. As for transitions to self-employment the positive effect on transitions to self-employment found in 1996 when the pair September/October is used as contrast is statistically significant.

	Difference in Differences in Probability			
	Informality		Self-employment	
Formal Sector	estimates	se	estimates	se
May-95	-0.0168	[0.015]	0.0004	[0.006]
Oct-95	0.0206	[0.015]	0.0005	[0.006]
May-96	-0.0230	[0.019]	-0.0003	[0.008]
Oct-96	-0.0213	[0.017]	0.0177	[0.007]**
May-97	-0.0141	[0.019]	0.0017	[0.008]
Oct-97	-0.0608	[0.021]**	-0.0015	[0.007]
May-98	0.0012	[0.022]	0.0024	[0.007]
Oct-98	0.0060	[0.021]	0.0064	[0.006]
May-99	0.0238	[0.020]	-0.0006	[0.008]
Oct-99	0.0043	[0.019]	-0.0118	[0.007]

*Significant at 5%, **Significant at 1%.

Table 5.13: Changes in Transition to Informality: Kernel Propensity Score Matching (1995-1999)

Matching: Nearest Neighbour based results

For the 1995-1999 period the NN estimates of the differences in the probability of transition to nonemployment for the pooled sample are positive. The only negative estimate is found when the pair of transitions September/October 1995 is used as contrast for the May 1995 increase. The positive estimates of changes in the probability of transition to nonemployment for 1996 (when the pair September/October is used as contrast) and for 1997, 1998 and 1999 (when the pair March/April is used

	Difference in Differences in Probability					
	Nonemployment		Unemployment		Inactivity	
Pooled Sample	estimates	se	estimates	se	estimates	se
May-95	0.001	[0.019]	-0.006	[0.009]	0.007	[0.017]
Oct-95	-0.021	[0.016]	-0.004	[0.009]	-0.017	[0.015]
May-96	0.021	[0.018]	0.010	[0.008]	0.011	[0.016]
Oct-96	0.041	[0.020]*	0.021	[0.010]*	0.021	[0.018]
May-97	0.049	[0.022]*	0.000	[0.010]	0.049	[0.020]**
Oct-97	0.006	[0.023]	-0.004	[0.010]	0.011	[0.021]
May-98	0.061	[0.021]*	0.012	[0.012]	0.049	[0.017]**
Oct-98	0.038	[0.020]	0.019	[0.009]*	0.018	[0.019]
May-99	0.054	[0.022]**	0.004	[0.011]	0.051	[0.020]**
Oct-99	0.010	[0.021]	0.010	[0.012]	0.000	[0.019]
Formal Sector						
May-95	-0.007	[0.017]	-0.002	[0.008]	-0.005	[0.015]
Oct-95	-0.007	[0.020]	0.005	[0.009]	-0.013	[0.018]
May-96	0.022	[0.020]	0.009	[0.009]	0.012	[0.018]
Oct-96	0.023	[0.023]	-0.003	[0.011]	0.026	[0.020]
May-97	0.009	[0.027]	-0.006	[0.013]	0.015	[0.023]
Oct-97	0.007	[0.025]	0.001	[0.013]	0.006	[0.022]
May-98	0.034	[0.023]	0.003	[0.011]	0.032	[0.022]
Oct-98	0.015	[0.016]	0.015	[0.016]	0.000	[0.023]
May-99	0.077	[0.027]**	0.013	[0.011]	0.063	[0.025]**
Oct-99	0.018	[0.025]	0.015	[0.014]	0.003	[0.021]
Informal Sector						
May-95	0.007	[0.033]	-0.013	[0.017]	0.020	[0.029]
Oct-95	-0.004	[0.030]	0.001	[0.017]	-0.004	[0.027]
May-96	0.025	[0.034]	0.022	[0.015]	0.003	[0.032]
Oct-96	0.001	[0.030]	0.023	[0.015]	-0.022	[0.027]
May-97	0.105	[0.044]**	0.025	[0.017]	0.079	[0.039]*
Oct-97	0.045	[0.046]	0.035	[0.029]	0.010	[0.038]
May-98	0.103	[0.035]**	0.028	[0.018]	0.074	[0.029]**
Oct-98	0.059	[0.037]	0.005	[0.017]	0.055	[0.034]
May-99	0.014	[0.040]	-0.023	[0.022]	0.037	[0.035]
Oct-99	0.003	[0.037]	-0.006	[0.024]	0.008	[0.030]

*Significant at 5%, **Significant at 1%.

Table 5.14: Changes in Transition to Non-Employment: Nearest Neighbour Propensity Score Matching (1995-1999)

as contrast) are statistically significant. The results for 1997, 1998 and 1999 seem to be driven by an increase in the probability of transition to inactivity, whereas the result for 1996 is due to the increase in transitions to unemployment.

As for the sample of formal sector workers, all the estimates of changes in the probability of moving to nonemployment are positive with the exception of the estimate for 1995 regardless the pair of months used as contrast. However, only for 1999 is the positive impact of the minimum wage hike on nonemployment statistically significant when the pair March/April is used as contrast. This result seems to be driven by transition to inactivity.

As for the sample of informal sector workers all the estimates of changes in the probability of moving to nonemployment are positive with the exception of the estimate for 1995 when the pair September/October is used as contrast. However, only for 1997 and 1998 when the pair March/April is used as contrast is the positive impact of the minimum wage hike on transitions to nonemployment significant. This result seems to be driven by transition to inactivity.

Table 5.15 shows that the differences in the probability in transition to informality also tend to show a positive sign. Only for the years 1996 and 1997 when the pair September/October is used as control are the estimates negative. However, the only significant effect is exactly the negative estimate for 1997. Transitions to self-employment, however, have a much more volatile pattern and in general the estimates display a negative sign. Nevertheless, the only significant estimate is the positive effect found for 1996 when the pair September/October is used as contrast[74].

	Difference in Differences in Probability			
	Informality		Self-employment	
Formal Sector	estimates	se	estimates	se
May-95	0.008	[0.018]	0.006	[0.008]
Oct-95	0.009	[0.019]	-0.004	[0.008]
May-96	0.008	[0.028]	0.006	[0.011]
Oct-96	-0.021	[0.025]	0.021	[0.010]*
May-97	0.033	[0.025]	-0.007	[0.012]
Oct-97	-0.069	[0.029]**	-0.007	[0.010]
May-98	0.040	[0.028]	-0.006	[0.011]
Oct-98	0.000	[0.027]	-0.003	[0.009]
May-99	0.031	[0.028]	0.003	[0.012]
Oct-99	0.056	[0.031]	-0.026	[0.013]*

*Significant at 5%, **Significant at 1%.

Table 5.15: Changes in Transition to Informality: Nearest Neighbour Propensity Score Matching (1995-1999)

5.6.3 Assessing Comparison Groups for the Non-parametric Results

A good comparison group should not display any strong effect after minimum wage increases. For instance, if workers marginally further up the wage distribution in relation to the minimum wage are good substitutes for minimum wage workers, one should expect their probability of moving to nonemployment to decrease, making the effect of the minimum wage even larger. Another possibility is that workers whose wages are indexed to the minimum wage would also be "at risk" of losing their jobs. If this were the case, the comparison group that includes multiples of minimum wage such as 1.5, 2, 2.5 or 3 times the minimum wage would also face a higher probability of moving to nonemployment after minimum wage hikes[75]. The real figure, however, must be a mixture of these two effects.

[74] This latter result is also found in the kernel strategy (at 5%) and in the probit estimates (at 10%).

[75] The wage indexation of the 1980's should have a similar effect. The lack of disemployment effect over this period may be due to shifts in the whole wage distribution caused by the indexation rules.

	POOLED SAMPLE					
	Nonemployment		Unemployment		Inactivity	
Kernel Estimates	estimates	se	estimates	se	estimates	se
May-82	-0.013	[0.006]*	-0.006	[0.003]	-0.007	[0.005]
Nov-82	-0.001	[0.005]	0.001	[0.004]	-0.002	[0.004]
May-83	-0.009	[0.005]	-0.011	[0.004]**	0.001	[0.004]
Nov-83	-0.008	[0.007]	-0.003	[0.005]	-0.006	[0.005]
May-84	0.016	[0.008]*	0.007	[0.006]	0.009	[0.005]
Nov-84	-0.006	[0.082]	-0.009	[0.007]	0.003	[0.004]
May-85	-0.010	[0.006]	-0.001	[0.003]	-0.009	[0.003]
Nov-85	-0.005	[0.009]	-0.009	[0.006]	0.004	[0.006]
May-95	0.012	[0.009]	0.004	[0.005]	0.008	[0.008]
Oct-95	0.002	[0.009]	-0.004	[0.005]	0.006	[0.008]
May-96	0.003	[0.009]	-0.002	[0.004]	0.005	[0.008]
Oct-96	-0.009	[0.010]	-0.004	[0.004]	-0.005	[0.009]
May-97	-0.009	[0.011]	0.008	[0.005]	-0.017	[0.011]
Oct-97	-0.007	[0.011]	-0.001	[0.006]	-0.005	[0.011]
May-98	-0.038	[0.011]**	-0.010	[0.005]*	-0.028	[0.009]**
Oct-98	-0.017	[0.009]	-0.007	[0.005]	-0.010	[0.006]
May-99	-0.028	[0.011]**	-0.003	[0.005]	-0.025	[0.010]**
Oct-99	0.001	[0.010]	-0.005	[0.005]	0.007	[0.009]
Nearest Neighbor						
May-82	-0.013	[0.008]	-0.010	[0.005]	-0.004	[0.007]
Nov-82	0.002	[0.006]	0.005	[0.004]	-0.003	[0.005]
May-83	0.000	[0.008]	-0.007	[0.006]	0.007	[0.006]
Nov-83	-0.009	[0.009]	-0.001	[0.007]	-0.008	[0.006]
May-84	0.018	[0.009]*	0.008	[0.006]	0.010	[0.082]
Nov-84	-0.013	[0.010]	-0.009	[0.009]	-0.004	[0.006]
May-85	-0.014	[0.007]*	-0.001	[0.004]	-0.013	[0.005]*
Nov-85	-0.002	[0.010]	-0.010	[0.007]	0.008	[0.007]
May-95	0.022	[0.014]	0.010	[0.007]	0.011	[0.012]
Oct-95	0.023	[0.013]	0.001	[0.008]	0.022	[0.011]
May-96	0.007	[0.014]	0.002	[0.006]	0.006	[0.013]
Oct-96	-0.013	[0.015]	-0.010	[0.009]	-0.003	[0.013]
May-97	-0.015	[0.016]	0.007	[0.008]	-0.022	[0.015]
Oct-97	-0.010	[0.016]	0.006	[0.009]	-0.016	[0.015]
May-98	-0.036	[0.016]*	-0.011	[0.009]	-0.025	[0.013]*
Oct-98	-0.017	[0.014]	-0.006	[0.007]	-0.011	[0.011]
May-99	-0.029	[0.016]	-0.007	[0.008]	-0.021	[0.014]
Oct-99	-0.016	[0.015]	*-0.016*	[0.008]***	-0.001	[0.013]

*Significant at 5%, **Significant at 1%, *** Significant at 10%.

Table 5.16: Difference in Probability Before and After the Treatment for the Control Group (Pooled Sample)

One way to assess how these effects jeopardize the quality of the comparison group is to estimate the differences in the probability of transition to nonemployment before and after the minimum wage increase for the comparison group sample. For the comparison group to be of good quality this difference must be small (close to zero) and not statistically significant. Tables 5.16 to 5.18 bring the results of these estimates for both kernel and nearest neighbour matching.

As for the Kernel estimates the results in Table 5.16 suggest that the negative effect in transitions to nonemployment found in May 1984 and the positive effect of the minimum wage on transitions to unemployment in May 1983 were due to statistically significant changes in the probability of transitions for the control group. The only robust result is the fall in transitions to self-employment for formal sector workers in May 1982. The results for the late 1990's are much more robust, since only the positive effect on transitions to nonemployment observed in 1998 does not pass the test. The positive effect on transitions to nonemployment found for 1996 and 1997 are robust to the test of quality of the control group.

As for the NN estimates the robustness checks results point to some effect of the minimum wage hikes on transitions to nonemployment and inactivity. The results for November 1983 and 1984 are robust for the pooled sample as well as the results for November 1983 and May 1985 for formal sector workers, and May 1983 and November 1984 for informal sector workers. The increase in transitions from the formal sector to self-employment in November 1983 and 1984 is not robust. For the period 1995-1999, the results for May 1998 using the pair March/April for the pooled sample and for informal sector workers did not pass the control group quality test as well as the result for May 1999 for formal sector workers when the pair March/April is used as contrast. All other results are robust to the quality of the comparison group.

	FORMAL SECTOR									
	Nonemployment		Unemployment		Inactivity		Informality		Self-employment	
Kernel Estimates	estimates	se	estimates	se	estimates	se	estimates	se	estimates	se
May-82	-0.010	[0.006]	-0.005	[0.004]	-0.005	[0.004]	-0.011	[0.006]*	0.003	[0.002]
Nov-82	0.005	[0.005]	0.002	[0.003]	0.002	[0.004]	-0.006	[0.005]	0.001	[0.002]
May-83	-0.010	[0.079]	-0.011	[0.005]*	0.001	[0.004]	0.003	[0.005]	-0.002	[0.003]
Nov-83	-0.007	[0.009]	-0.002	[0.006]	-0.005	[0.082]	-0.003	[0.007]	-0.003	[0.002]
May-84	0.005	[0.009]	-0.001	[0.005]	0.006	[0.007]	-0.005	[0.010]	0.007	[0.010]
Nov-84	-0.001	[0.007]	-0.007	[0.006]	0.006	[0.005]	0.015	[0.007]*	0.002	[0.002]
May-85	-0.008	[0.006]	-0.003	[0.004]	-0.004	[0.005]	0.007	[0.005]	0.003	[0.002]
Nov-85	0.001	[0.008]	-0.002	[0.005]	0.003	[0.005]	0.037	[0.010]**	-0.010	[0.005]*
May-95	0.006	[0.008]	0.003	[0.003]	0.003	[0.007]	0.021	[0.010]*	-0.001	[0.005]
Oct-95	-0.003	[0.009]	-0.004	[0.005]	0.001	[0.008]	0.004	[0.011]	0.001	[0.003]
May-96	0.003	[0.007]	-0.001	[0.004]	0.004	[0.007]	0.019	[0.012]	0.003	[0.004]
Oct-96	-0.014	[0.011]	-0.007	[0.006]	-0.008	[0.009]	0.028	[0.011]**	-0.004	[0.005]
May-97	0.012	[0.014]	0.010	[0.007]	0.002	[0.013]	0.017	[0.010]	-0.002	[0.006]
Oct-97	0.020	[0.016]	0.003	[0.009]	0.017	[0.014]	0.019	[0.013]	0.000	[0.005]
May-98	-0.023	[0.013]	-0.003	[0.007]	-0.020	[0.011]	-0.006	[0.013]	-0.004	[0.005]
Oct-98	-0.013	[0.013]	-0.008	[0.011]	-0.005	[0.008]	0.011	[0.010]	-0.001	[0.005]
May-99	-0.016	[0.012]	-0.002	[0.005]	-0.015	[0.011]	-0.002	[0.011]	-0.003	[0.006]
Oct-99	-0.009	[0.012]	-0.008	[0.007]	-0.001	[0.009]	0.009	[0.010]	0.006	[0.005]
Nearest Neighbor										
May-82	-0.003	[0.008]	-0.006	[0.005]	0.003	[0.007]	-0.005	[0.007]	0.001	[0.002]
Nov-82	0.005	[0.007]	-0.001	[0.005]	0.006	[0.005]	-0.003	[0.007]	-0.003	[0.003]
May-83	0.001	[0.009]	-0.005	[0.007]	0.005	[0.006]	0.012	[0.006]	-0.005	[0.003]
Nov-83	-0.010	[0.010]	0.001	[0.007]	-0.011	[0.007]	0.006	[0.008]	-0.006	[0.003]**
May-84	0.009	[0.009]	-0.006	[0.005]	0.014	[0.008]	-0.005	[0.010]	0.008	[0.010]
Nov-84	-0.008	[0.010]	-0.012	[0.008]	0.003	[0.005]	0.009	[0.007]	0.001	[0.003]
May-85	-0.015	[0.008]	-0.006	[0.005]	-0.009	[0.006]	0.008	[0.007]	0.002	[0.003]
Nov-85	-0.008	[0.009]	-0.006	[0.007]	-0.003	[0.007]	0.037	[0.010]**	-0.010	[0.005]*
May-95	0.015	[0.013]	0.007	[0.006]	0.007	[0.012]	0.017	[0.013]	-0.001	[0.007]
Oct-95	0.003	[0.014]	-0.003	[0.007]	0.006	[0.013]	0.009	[0.014]	0.004	[0.006]
May-96	-0.007	[0.015]	0.009	[0.009]	0.012	[0.018]	0.021	[0.018]	0.001	[0.008]
Oct-96	-0.013	[0.017]	0.004	[0.009]	-0.017	[0.015]	0.003	[0.016]	-0.007	[0.009]
May-97	0.032	[0.019]	0.015	[0.010]	0.017	[0.018]	-0.002	[0.017]	0.006	[0.010]
Oct-97	0.014	[0.020]	0.000	[0.011]	0.014	[0.017]	0.032	[0.021]	-0.004	[0.007]
May-98	-0.015	[0.016]	0.000	[0.009]	-0.015	[0.014]	-0.013	[0.019]	0.003	[0.010]
Oct-98	-0.006	[0.019]	-0.004	[0.014]	-0.003	[0.014]	0.026	[0.020]	0.010	[0.007]
May-99	-0.056	[0.020]**	-0.012	[0.007]	-0.044	[0.019]*	-0.015	[0.020]	-0.001	[0.010]
Oct-99	-0.011	[0.018]	-0.012	[0.011]	0.002	[0.014]	-0.018	[0.020]	0.006	[0.010]

*Significant at 5%, **Significant at 1%.

Table 5.17: Difference in Probability Before and After the Treatment for Control Group (Formal Sector)

Overall a closer look at Table 5.19 suggests that the minimum wage hikes have provoked a larger disemployment effect in the late 1990's than in the early 1980's. This effect is stronger for the pooled sample, than for any of the separate samples we have investigated. Transitions to the informal sector or to self-employment were rarely significant for both sample periods. Actually, we find some evidence for the late 1990's that points to a decrease in transitions to both informal sector and self-employment after a minimum wage hike.

It seems that assessing the quality of the control group is a good way to avoid misleading results when evaluating the effect of minimum wage increase on employment transitions. From this section, we could conclude that several of our difference-in-difference matching estimates that seemed to lead to a positive effect of the minimum wage on transitions to nonemployment were due to significant changes in the transition probabilities of the comparison group. Whether this is due to spill-over effects or different reaction between treated and comparison group to simultaneous macroeconomic shocks is not clear. More research is needed to disentangle these two effects.

5.7 Conclusion

This chapter aimed to evaluate the effect of several episodical minimum wage increases on employment transitions in the early 1980's and late 1990's in Brazil. We evaluate this effect for the pooled sample of formal and informal workers and for each group separately. We emphasize the need to incorporate the informal sector in the analysis for two reasons. First, due to possible dynamics effects of minimum wage hikes on the wage of informal sector workers as highlighted in Welch (1974) and Mincer (1976) models. Second due to the well-documented increase in the importance of the minimum wage in the determination of wage or wage increases for informal sector workers. Thus minimum wage hikes are quite likely to affect the employment transition of informal sector workers either because of (indirect) dynamic effects or because of the informal indexation of the informal sector wages to the minimum wage.

We estimate the effect of the minimum wage hike using difference-in-differences in a parametric way (via probit) and in a non-parametric way (via propensity score matching). To define treated and control groups in the baseline period that was used as contrast (in the difference-in-differences) for the actual minimum wage increase we defined pseudo-experiments both before (in the previous month) and after the

	INFORMAL SECTOR					
	Nonemployment		Unemployment		Inactivity	
Kernel Estimates	estimates	se	estimates	se	estimates	se
May-82	-0.001	[0.016]	0.008	[0.010]	-0.009	[0.013]
Nov-82	-0.033	[0.020]	-0.005	[0.011]	-0.028	[0.017]
May-83	-0.031	[0.019]	-0.026	[0.015]	-0.004	[0.015]
Nov-83	-0.039	[0.025]	-0.002	[0.019]	-0.037	[0.018]*
May-84	0.061	[0.026]**	0.044	[0.020]*	0.017	[0.021]
Nov-84	-0.008	[0.024]	0.018	[0.014]	-0.026	[0.017]
May-85	-0.009	[0.023]	0.012	[0.010]	-0.021	[0.021]
Nov-85	0.015	[0.018]	-0.010	[0.011]	0.025	[0.014]
May-95	0.004	[0.023]	-0.003	[0.011]	0.007	[0.020]
Oct-95	0.022	[0.020]	0.000	[0.009]	0.023	[0.018]
May-96	-0.015	[0.023]	-0.007	[0.007]	-0.008	[0.022]
Oct-96	0.018	[0.019]	-0.003	[0.006]	0.021	[0.018]
May-97	-0.033	[0.020]	0.002	[0.009]	-0.035	[0.018]
Oct-97	-0.034	[0.024]	-0.018	[0.013]	-0.015	[0.021]
May-98	-0.049	[0.017]**	-0.016	[0.008]	-0.033	[0.015]*
Oct-98	-0.032	[0.017]	-0.002	[0.008]	-0.031	[0.014]*
May-99	-0.032	[0.021]	-0.004	[0.012]	-0.028	[0.019]
Oct-99	-0.005	[0.020]	0.000	[0.010]	-0.005	[0.018]
Nearest Neighbor						
May-82	-0.016	[0.020]	0.007	[0.012]	-0.227	[0.017]
Nov-82	-0.026	[0.025]	0.101	[0.081]	-0.036	[0.021]
May-83	-0.032	[0.022]	-0.020	[0.016]	-0.013	[0.017]
Nov-83	-0.027	[0.027]	0.026	[0.020]	-0.053	[0.019]**
May-84	0.064	[0.028]*	0.041	[0.020]*	0.023	[0.023]
Nov-84	-0.009	[0.019]	0.014	[0.012]	-0.023	[0.013]
May-85	-0.003	[0.025]	0.030	[0.012]	-0.033	[0.023]
Nov-85	0.014	[0.019]	-0.008	[0.012]	0.022	[0.014]
May-95	-0.001	[0.027]	0.007	[0.014]	-0.008	[0.024]
Oct-95	0.021	[0.025]	-0.008	[0.012]	0.029	[0.022]
May-96	0.004	[0.028]	-0.009	[0.010]	0.013	[0.027]
Oct-96	0.021	[0.026]	-0.008	[0.010]	0.028	[0.023]
May-97	-0.044	[0.032]	-0.010	[0.014]	-0.034	[0.029]
Oct-97	-0.069	[0.034]*	-0.054	[0.024]*	-0.015	[0.025]
May-98	-0.057	[0.024]*	-0.014	[0.014]	-0.043	[0.020]*
Oct-98	-0.012	[0.024]	0.012	[0.011]	-0.024	[0.022]
May-99	0.001	[0.031]	0.011	[0.018]	-0.010	[0.027]
Oct-99	-0.004	[0.029]	-0.006	[0.017]	0.001	[0.028]

*Significant at 5%, **Significant at 1%.

Table 5.18: Difference in Probability Before and After the Treatment for Control Group (Informal Sector)

	Nonemployment	Unemployment	Inactivity	Informality	Self-Employment
Kernel					
Pooled Sample					
Oct-96	YES				
May-97	YES				
Oct-98	YES				
Formal Sector					
May-82					YES
Oct-96	YES				YES
Oct-97				YES (-)	
Informal Sector					
Oct-98	YES				
NN					
Pooled Sample					
Nov-83	YES		YES		
Nov-84	YES		YES		
Oct-96	YES	YES			
May-97	YES		YES		
Oct-98	YES	YES			
May-99	YES		YES		
Formal Sector					
Nov-83	YES		YES		
May-85	YES				
Oct-96					YES
Oct-97				YES(-)	
Oct-99					YES(-)
Informal Sector					
May-83	YES	YES			
Nov-84	YES		YES		
May-97	YES		YES		

(-) Means that there was a decrease in transitions.

Table 5.19: Summary of the Effect of the Minimum Wage Hikes

minimum wage increase[76] (5 months later).

For the pooled sample we only find evidence of disemployment effects for some of the episodical increases of the late 1990's. According to the probit estimates this occurred in 1996 and 1998 when the pair September/October is used as control. The kernel matching indicates that the increase in transitions to nonemployment was significant in 1996 for the pair September/October, in 1997 for the pair March/April and in 1998 for the pair September/October[77]. The NN estimates are in line with the kernel estimates, but it also shows a statistically significant increase in 1999 when the pair March/April is used as contrast.

As for the formal sector, the results are in line with the probit estimates, there is no disemployment effect on the early 1980's regardless of the method used. For the late 1990's only the kernel matching point to a positive and significant effect in 1996 when the pair September/October is used as contrast. We also find no strong evidence that minimum wage hikes led to transitions to informality or to self-employment.

[76]Notice that this second pseudo-experiment was only possible for the sample of the late 1990's.

[77]The result for 1998 using the pair March/April as contrast is not robust to the quality of control group.

As for the informal sector, the probit estimates are quite similar to the estimates for the pooled sample, indicating disemployment effects in 1996 and 1998 and no effect on the early 1980's. The kernel matching only shows significant disemployment effects in 1998 (when the pair September/October is used) whereas the NN matching indicates significant disemployment effects in 1997 (when the pair March/April is used) and in May 1983.

All in all it seems that the probit estimates and the two propensity score matching do not diverge very much in their estimates, this is particularly true for the kernel estimates. The results indicate that the minimum wage hikes were more likely to have disemployment effects in the late 1990's. These results suggest that the wage indexation of the early 1980's avoided the disemployment effect of the minimum wage as long as it affected the whole wage distribution[78]. However, one puzzling result refers to the minimum wage increase of May 1995. The increase in the real value of the minimum wage observed in that date should make disemployment effects more likely to be found in that episode. However, all estimates point to a negative (but statistically insignificant) effect on transitions to nonemployment for that year. One possible explanation for this result is that the minimum wage was so low before that increase that it was not binding for formal and informal workers.

[78]More research is needed to establish a causal link between these two evidences.

Chapter 6

Conclusion

Informal labour markets in developing countries have always been a grey area between development economics and labour economics. In his seminal paper Fields (1975) emphasizes their transitory nature. The informal sector is viewed as a waiting stage for rural migrants into the urban labour market in the tradition of the development economics literature. Unlike Fields, ILO's (1972) approach stresses the structural characteristics of the informal sector which is conceptualized as the set of firms at the lower-end of the size distribution. In this regard, the informal sector is a permanent feature of developing economies and the main focus of public policies aimed at alleviating poverty and fostering economic growth. This thesis departed from these approaches and opted to treated the informal sector in the tradition of the labour economics literature, which concentrates on the wage differential between similar work in both sectors and on how to measure segmentation. For this reason, our definition of informal workers included only non-registered employees. Our view is that this sector cannot be mixed with the self-employed and small firms owners as it is commonly found in the literature. Therefore, we focus on the labour economics literature to tackle the stilysed facts under scrutiny on this thesis.

This thesis was motivated by the empirical evidence that the informal sector has a huge size in the Brazilian economy and that its size has increased after the market-oriented reforms. Moreover, other changes were in motion: the wage gap between workers in the formal and informal sector decreased after the reforms and the minimum wage became much more important for informal than formal workers. Therefore, we had the opportunity to discuss two lines of research. The first one, the existence of segmentation between formal and informal workers, brings the job queue literature to the debate on how to measure the wage differential between otherwise similar workers who are allocated in different sectors. We were particularly disturbed by the previous findings in the literature that claimed that informal sector workers in Brazil allocated themselves into the sector where they have comparative advantage, while household surveys revealed that the great majority of informal sector workers would like to switch to formal sector jobs. Economists do not believe what people say, but rather in what people do. However, in order to uncover why and how people do what they do in the absense of a credible counterfactual, what people say that they would like to do can be a good guide when deciding which methodology to apply to tackle a problem. This is particularly the case when the researcher does not observe"how people act". This is the case of the severe partial observability that constitutes one of the main characteristics of the sector allocation process in the presence of job queue. We only observe a worker in the formal sector after he/she queued for a formal job and was chosen from it. Therefore, we do not observe the "in the queue" status. The use of subjective question about the desire on changing to a formal job, can help to reduce this severe partial observability and provide a better understanding of the sector allocation process as showed in Chapter 3.

The second line refers to the use of policy changes to understand the dynamics of the informal sector. In the absence of labour legislation reforms, the other market-oriented reforms of the early 1990's and the minimum wage policy in Brazil are great opportunities to investigate how the informal sector reacts to policy changes. The trade liberalisation programme of the early 1990's is thought to have reduced the bargaining power of formal sector workers, helping to reduce the wage gap between formal and informal workers due to the fall in rents. Similarly, common sense has blamed trade liberalisation for the fall in the proportion of registered workers. We investigated this hypothesis in Chapter 4 of this thesis. The minimum wage hikes was always thought to positively affect the size of the informal sector. In chapter 5 we analyse how the minimum wage policy has affected employment transitions and whether this view as a source of informality is borne by the data. In the following paragraphs we summarise the main findings of this thesis, its caveats, future research lines and discuss public policies implications of some findings.

In Chapter 2, we quickly reviewed the literature on segmentation in developing countries and explained our option to use only non-registered workers as informal ones. Besides, we established several stylised facts about the informal sector so that we could set the scene for the other chapters. Here is a summary of these facts. The informal sector made up 40% of private sector employees in 1999. That proportion has increased during the 1990's, and the bulk of the increase has coincided with the market-oriented reforms. Alongside the increase in the proportion of informal workers, a fall in the wage gap between formal and informal sector workers was also witnessed. The fall in the wage gap contributed for the slight decrease in inequality between 1981-1999. According to the results of the Field's (2002) decomposition, the variable "possession of a work-card" was the second most important factor behind the reduction in wage inequality after education. The decomposition of the changes in the wage differential between formal and informal workers revealed that (the quantity of) unobservables was the main factor responsible for the reduction. Formal and informal workers became more similar in relation to their observables, and to their returns to productive characteristics, but it was mostly due to changes in unobservables that the wage gap narrowed. The Juhn-Murphy-Pierce "full-sample distribution accounting scheme" also revealed that it was the reduction in inequality at the bottom of the distribution that led to an overall reduction in inequality. Between 1981 and 1999 there was a sharp fall in the non-compliance with the minimum wage legislation among both formal and informal sector workers. Furthermore, the minimum wage from 1995 onwards provoked a much more marked spike in the wage distribution for informal sector than for formal sector workers. On the top of that, there were more informal sector workers receiving wage increases equal to the minimum wage hike than formal sector workers in the late 1990's.

In chapter 3, we investigated the question of segmentation, that we had briefly discussed in the first part of Chapter 2, in the lines of the formal and informal sector separation. We used a job queue approach in order to establish the existence of segmentation. Exploiting a subjective question on the desire of informal sector workers to switch to a formal job, we were able to test the existence of a queue for formal sector job, to estimate the length of the queue for different type of workers, and to correct the wage equation for formal and informal workers using two-selectivity criteria, based on the the probability of being in the queue and on the probability of being chosen from the queue. Additionally, we estimated a structural model to assess the role of the wage gap in the determining the "in the queue" status.

Methodologically, we assessed the reliability of the Abowd-Farber bivariate probit as well as Poirier bivariate probit in dealing with partial observability models such as the one developed in that chapter. Our results suggest that the use of the question on the desire to join the formal sector reduces the severe partial observability on the "in the queue" status and yields results much less sensitive to small changes in specifications. All approaches used in chapter 3 point to the existence of a job queue for formal jobs. Moreover, previous informal sector workers, "new entrants", illiterate and non-white workers are the groups for whom the length of the queue is larger. If one considers that queueing for formal jobs in the informal sector is a second best option that can jeopardize the individual chances in the labour market, then those groups of workers should be the focus of public policies that aim at facilitating their access to formal (registered) jobs. We also showed that the use of a univariate process to depict the sector allocation along the lines of the formal and informal sector hides interesting features of this process. For instance, whereas education has almost no impact on the decision to queue for formal jobs, it is the main variable determining the decision to hire a worker. As for the structural model, the results suggest that the wage gap is the main variable in determining the decision to join the queue. A future research agenda on this field would be to relax the parametric assumptions about the distribution of the residuals of the equations and estimate distribution-free models for the "in the queue" and the "chosen from the queue" equation and then estimate semi-parametric corrected wage equations.

In chapter 4, we looked carefully at the coincidence between the fall in the wage gap between formal and informal workers and the increase in the proportion of informal sector workers and the trade liberalisation process. In principle, the trade liberalisation process could have led to adjustment in the productive structure of the Brazilian economy so that small firms and sectors more likely to employ workers "informally" would benefit from the liberalization, whereas formal-prone firms would be forced to downsize their work-force in order to boost productivity. In this case, trade liberalisation would have caused a fall in the proportion of registered (formal) workers. Similarly, if the more competitive environment brought about by the trade liberalisation hurt larger firms that enjoyed rents due to the lack of competition, and shared these rents with their employees, then the fall in the wage gap would be a possible outcome of trade reforms.

In order to assess these hypotheses we put forward three reduced form approaches. First, we exploited industry variation of trade-related measures: effective tariffs, import penetration and export orientation ratios for a panel of 17 tradable manufacturing industries. The results for the fixed-effect models suggest that the increase in the import penetration ratio had a negative effect on the wage premium for registered workers in the manufacturing sector. As for the increase in the proportion of registered workers, we do not find any robust evidence that trade liberalisation could be behind this phenomenon.

Second, we exploited regional variation of industry dispersion within the country, so that we were able to test the impact, if any, of trade liberalisation on the entire labour market and not only in the manufacturing sector (spillover effect). We did not find any evidence of spillover effect in the case of the fall in the wage gap. Moreover, the evidence that the increase in the import penetration ratio led to a fall in the proportion of registered workers in the entire labour market was not robust to the use of lagged regressors. Therefore, the overall narrowing in the wage gap between formal and informal sector workers did not seem to be related to the trade liberalisation.

Third, we added the non-tradable sector to the industries used in the first strategy and built pseudo-cohorts from a repeated cross-section data set. This latter procedure allowed us to use the non-tradable sector as a control group for the directly affected tradable sector. Moreover, the use of pseudo-cohorts allowed us to control for industry-cohort fixed-effects. The results showed that both the fall in tariffs as measured by the *openness* index and the increase in import penetration ratio as measured by the *closedness* index led to a fall in the wage gap between formal and informal workers. As for the proportion of registered workers, the fixed effect models suggest that the fall in tariffs led to a higher proportion of registered workers, whereas the increase in the import penetration ratio led to a fall in the proportion of registered workers. This result is puzzling and given the results of the first strategy, we concluded that the findings for the wage gap are much more robust than this last mixed result for the proportion of registered workers.

A possible caveat of the approach adopted in chapter 4 is that we do not look at the effects of trade liberalisation on the worker's mobility pattern. Focusing on the effect on the proportion of registered workers and on the wage differential give us only a partial account of the possible effects of those changes in the labour market. Another interesting topic of research would be to analyse what happened to the size of the queue for formal jobs after the trade liberalisation. Unfortunately, we do not have surveys that contain a question similar to the one used in Chapter 3 to identify informal workers in the queue for formal jobs for periods after the trade liberalisation. However, as the wage differential is one of the main determinants of the queue for formal jobs, it is reasonable to infer that the size of the queue may have diminished after trade liberalisation due to the narrowing of the gap between formal and informal workers.

In Chapter 5 we tackled two issues related to the informal sector and minimum wage policy. First we documented the changes in the way the wage distribution for formal and informal sector workers was shaped by the minimum wage and how minimum wage hikes related to contemporaneous increases in the average wage of each percentile for formal and informal workers. We showed that there was a clear loss of importance of the minimum wage for formal sector workers simultaneously with an increase of its importance for informal sector workers from the 1980's to the 1990's. Therefore, the minimum wage policy in Brazil is likely to have a direct effect on informal sector workers, at least, in the second half of the 1990's, besides the indirect affect pointed out in models like the ones developed by Welch (1974) and Mincer (1976). Second, we investigated the effect of the minimum wage hikes on employment transitions in Brazil for the pooled sample of formal and informal workers and for each of them separately. Workers affected by minimum wage increases were compared with similar workers further up the wage distribution. To control for heterogeneity between the treated minimum wage workers and the control groups we used a difference-in-differences approach that compared treated and control groups in periods with nominal increases in the minimum wage with periods with no increase. In this last case the control and treated groups were defined as if there had been an increase in the minimum wage (pseudo-experiment)[1]. The main findings were that disemployment effects are much more likely to be found in the late 1990's (period without official indexation) than in the early 1980's. We did not find strong evidence of transitions to either informality or self-employment after minimum wage hikes. Methodologically, probit and propensity score matching do not differ much in their estimates.

One caveat of this procedure is the short term of the analysis. We analyse the transitions just one month after the minimum wage increase. The structure of the panel data, the 6-month interval increase of the minimum wage in the early 1980's and the high turnover rate in Brazil make very difficult any attempt of meaningfully estimating transitions based on longer periods. In terms of future research, one interesting topic would be to clarify why there was almost no disemployment effect in the late 1980's, whereas during the 1990's some of episodes show disemployment effect. One prime suspect would be the indexation rules of the 1980's, since they would have affected the whole wage distribution and were based in a formula that allowed firms to forecast with good precision the minimum wage hike.

In terms of public policy, the results of this thesis do not support the view that non-registered workers are satisfied with their status and have chosen to be non-registered because they have comparative advantage in that sector. Chapter 3 shows that this is absolutely not the case. Some researchers have claimed that the non-registered status may be an option for most workers who are working in the informal sector, this is not borne by the data, though. Camargo (2004), for instance, claims that there is a myth that informality is synonymous with lack of social protection. Of course, there is in Brazil a safety net not related to the labour market, but the access to several public funds are conditional on having a registered job and this does make a difference when queueing for a formal job.

[1]This strategy was applied in a parametric way via probit estimates for a variety of control groups and also in a nonparametric way using different propensity score matching methods.

Labour relation reforms should not seek to see the informal sector as an ideal type, nevertheless, it should curb legislation and practices that allow the lucky ones - the ones who get a formal job (chosen from the queue) - to appropriate rents and by this means perpetuate income inequality[2]. Sometimes, policy changes that aimed to affect other macroeconomic variables have spillover effects on the labour market. We saw in Chapter 4 that the trade liberalisation process has curbed the rents that accrued to manufacturing formal sector workers, and hence has reduced the wage gap between formal and informal sector workers. Similarly, minimum wage hikes, nowadays, have more effect on informal sector workers than on their formal counterparts. However, this sort of income policy has to be applied carefully, as it seems that the negative effect on employment transitions - mainly for informal sector workers - are more prevalent nowadays with the end of wage indexation than before during the wage indexation era. There is no easy way out to tackle the segmentation between registered and non-registered workers, but a better understanding of the way both groups have changed in the last two decades may shed some light on the way the government can change legislation in order to alleviate income inequality and eliminate labour market inefficiencies.

[2]It is worth to remember that former informal sector workers are the ones with the lowest probability to be chosen from the queue for formal jobs.

Chapter 7

References

Abowd, J. M. and Farber, H. S (1982) "Job Queues and the Union Status of Workers". *Industrial and Labor Relations Review*, Vol.35, N.3, pp. 354-367.

Abowd, J. M., Kramarz, F., Margolis, D. N., and Philippon, T. (2000) "The Tail of Two Countries". mimeo.

Ahn, S. C.(1992) "The Lagrangian multiplier test for a model with two selectivity criteria". *Economics Letters*, Vol.38, pp.9-15.

Arbache, J. S. (1999), "Trade Liberalisation and Inside Power: The Case of Brazil". Mimeo.

Arbache, J. S., Dickerson, A., Green, F. (2004) "Trade Liberalisation and Wages in Developing Countries". *The Economic Journal*, vol. 114, F73-F96.

Arbache, J. S., Corseuil, C. H. (2003) "Liberalizacao Comercial e Estrutura de Emprego e Salarios". Mimeo.

Barros, R. P., Camargo, J. M. and Sedlasek, G. (1992) "Os Tres Mercados: Segmentacao, Mobilidade e Desigualdade". IPEA. Mimeo.

Barros, R. P, Mello, R. and Pero, V. (1993) "Informal Labor Contract: A Solution or a Problem?". Texto para Discussao No. 291. IPEA.

Barros, R. P., Camargo, J. M. and Firpo, S. (1998) "Uma Avaliacao da Evolucao Recente do Diferencial de Rendimentos Formal/Informal. Ministerio do Trabalho". Mimeo.

Behrman, Jere. R, (1999) Labor Markets in Developing Countries In *Handbook of Labor Economics*, Vol.3, Ed. O. Ashenfelter and D. Card. Elsevier Science.

Behrman, Jere R., Birdsall, N. and Szekely, M. (2001) "Economic Policy and Wage Differential in Latin America" Penn Institute for Economic Research, Working Paper 01-048.

Berman, E. and Machin, S. (2000). "Skill-Biased Technology Transfer: Evidence of Factor Biased Technological Change in Developing Countries", mimeo.

Beyer, H.; Rojas, P. and Vergara, R. (1999), 'Trade liberalization and wage inequality', *Journal of Development Economics*, Vol.59 pp.103-123.

Bhaskar, V. and To, T. (1996) "Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition". Mimeo.

Blau, D. M. (1985) "Self-Employment and Self-Selection in Developing Country Labor Markets." *Southern Economic Journal.* Vol. 51 N. 2 pp. 351-363.

Blundell, R. and Dias, M. C. (2000)"Evaluation Methods for Non-experimental Data". *Fiscal Studies(2000).* Vol. 21, N.4, pp. 427-468.

Blundell, R., Dias, M. C., Meghir, C., and Van Reenen, J. (2001) "Evaluating the Employment Impact of a Mandatory Job Search Assistance Program". *IFS Working Paper* 01/20.

Borjas, George J. and Ramey, Valerie A. (1995) Foreign Competition, Market Power, and Wage Inequality. *Quarterly Journal of Economics.* Vol.110, pp.1075-1110.

Brown, Charles, Curtis Gilroy and Andrew Kohen (1982) "The Effect of the Minimum Wage on Employment and Unemployment". *Journal of Economic Literature*, Vol. 20, pp.487-582.

Brown, Charles (1999) "Minimum Wages, Employment and the Distribution of Income." In *Handbook of Labor Economics 3* pp. 2101-163, ed. Ashenfelter and Card.

Camargo, J. M. and Ramos, C. A. (1988). "A Revoluo Indesejada". Rio de Janeiro: Campus.

Camargo, J. M. (2004) "O Investimento como Producao: a logica da reforma trabalhista" In *Reformas no Brasil: Balanco e Agenda* pp. 339-358.

Card, David and Krueger, A. (1995) *Myth and Measurement: the New Economics of the Minimum Wage.* Princeton: Princeton University Press.

Carneiro, Francisco G., (2000) "Time Series Evidence on the Employment Effect of Minimum Wages in Brazil". Mimeo. http://ssrn.com/abstract=231875

Carneiro, F. G. and Henley, A. (2001) "Modelling Formal vs. Informal employment and earnings: microeconometric evidence for Brazil", *Annals of the XXIX National Meeting of Economics of ANPEC.*

Co, C., Gang, Ira N. and Yun, M. (2002) "Self-Employment and wage earing in Hungary during transition. Mimeo.

Corseuil, C. H. and Morgado, W. (2002) "Salrio Mnimo, Distribuio de Salrios e Emprego no Brasil". Mimeo.

Currie, J. and Harrison, A. (1997) "Sharing Costs: The Impact of Trade Reform on Capital and Labor in Marroco". *Journal of Labor Economics*, 15: s44-s71.

Currie, J. and Fallick, B. C. (1996) "The Minimum Wage and the Employment of Youth: Evidence form the NLSY". *The Journal of Human Resources.* Vol. 31, N. 2, pp.404-428.

Davis, D. (1996) "Trade Liberalization and Income Distribution". NBER Working Paper 5693.

Deaton, A. (1985) "'Panel Data From Time Series of Cross-Sections" *Journal of Econometrics*, Vol. 30, pp. 109-126.

Dickens, William T. and Lang, Kevin (1985) "A Test for Dual Labor Markets. *American Economic Review*, Vol. 75 N.4, pp. 792-805.

Dickerson, A., Green, F. and Arbache, J. S. (2001) "Trade Liberalization and the Returns to Education: A Pseudo-Panel Approach. Mimeo.

Doeringer and Piore (1971) *Internal Labor Markets and Manpower Analysis.* Ed. Sharpe.

Driffill, J., Meschi, M., and Ulph, A., (1998) "Product Market Integration and Wages: Evidence from a Cross-section of Manufacturing Establishments in the United Kingdom. *CEPR Discussion Paper* No. 1818.

Esfahani, Hadi S. and Salehi-Isfahani, D. (1989) "Effort Observability and Worker Productivity: Towards an Explanation of Economic Dualism", *The Economic Journal* Vol. 99, N.397, pp. 818-836.

Farber, H. S. (1983) The determination of the union status of workers". *Econometrica*, Vol. 51, pp. 1417-1437.

Fernandes, Reynaldo and Menezes-Filho, Naercio A. (2000) "A Evolucao da Desigualdade no Brasil Metropolitano entre 1983 e 1997", *Estudos Economicos* Vol. 30, N.4, pp. 549-569.

Fields, Gary S. (1975) "Rural/urban Migration, Urban Unemployment and Underemployment, and Job Search Activity in LACS", *Journal of Development Economics* Vol.2 pp.165-187.

Fields, Gary S. (2002) "Accounting for the Income Inequality and its Change: A New Method, with Application to the Distribution of Earnings in the United States." *Research in Labor Economics.* Forthcoming.

Foguel, M. N. (1998) "Uma Avaliao dos Efeitos do Salrio Mnimo sobre o Mercado de Trabalho no Brasil" *Texto para Discussao* IPEA N. 564.

Foguel, M. N., Corseuil, C. H., Barros, R. P., and Leite, P. G. (2000) "Uma Avaliao dos Impactos do Salrio Mnimo sobre o Nvel de Pobreza Metropolitana no Brasil" *Texto para discussao* N.739.

Fonseca, R. et al. (2000) "A Orientacao externa da industria de transformacao brasileira apos a liberalizacao commercial" *Revista de Economia Politica* 20(3) pp. 22-38.

Fortin et al. (1997) "Taxation, Wage Controls and the Informal Sector", *Journal of Public Economics* Vol.66 pp.293-312.

Funkhouser, E. (1997) "Demand-Side and Supply-Side Explanations for Barriers to Labor Market Mobility in Developing Countries: The Case of Guatemala". *Economic Development and Cultural Change.* Vol.45, N.2, pp. 341-366.

Galiani, S. and Sanguinetti, P. (2001) "Wage Inequality and Trade Liberalization: Evidence from Argentina". *Universidad Torcuato Di Tella,* mimeo.

Gindling, T. H. (1991) "Labor Market Segmentation and Determination of Wages in the Public, Private-Formal and Informal Sector in San Jose, Costa Rica." *Economic Development and Cultural Change.* Vol. 39 N.3 pp. 585-605.

Gindling, T. H. and Robbins, D. (2001) "Patterns and Sources of Changing Wage Inequality in Chile and Costa Rica During Structural Adjustment". *World Development* Vol.29, No. 4, pp. 725-745.

Goldberg, P. K. and Pavcnik, N. (2003) "The response of the informal sector to trade liberalisation". *NBER Discussion Paper* No. 9443.

Goldin, C and Katz, L. F. (1998) "The origins of technology-skill complementarity", *Quarterly Journal of Economics* Vol.113, pp. 693-732.

Gong, X, Van Soest, A. and Gomez, E. V. (2000) "Mobility in the Urban Labor Market: A Panel Data Analysis for Mexico. *IZA Discussion Paper* No. 213.

Gong, X. and Van Soest, A. (2002) "Wage Differentials and Mobility in the Labour Market: A Panel Data Analysis for Mexico". *Labour Economics* Vol.9 pp. 513-529.

Gonzaga, G., Menezes-Filho, N. and Terra, M. C. (2002a) "Earnings Differentials and Trade Liberalization in Brazil". Fundacão Getúlio Vargas, mimeo.

Gonzaga, G., Menezes-Filho, N. A., Camargo, J. M. (2002b) "Os Efeitos da Reducao da Jornada de Trabalho de 48 para 44 Horas Semanais em 1988. Miemo.

Green, F., Dickerson, A., Arbache, J. S. (2001) "A Picture of Wage Inequality and the Allocation of Labor through a Period of Trade Liberalization: The Case of Brasil". *World Development,* vol. 29, pp.1923-39.

Hay, Donald (1998) "The Post 1990 Brazilian Trade Liberalisation and Performance of Large Manufacturing Firms: Productivity, Market Share and Profits". *The Economic Journal* Vol.111 No. 473, pp. 620-641

Johnson, G., Stafford, F. (1999) "The Labor Market Implications of International Trade". In *Handbook of Labor Economics,* Vol.3. Ed. O. Ashenfelter and D. Card. Elsevier Science B. B.

Juhn, C., Murphy, K.M., and Pierce, B. (1991) "Accounting for the Slowdown in Black-White Wage Convergence". In *Workers and their Wages* Editor Marvin H. Kosters. The AEI Press.

Juhn, C., Murphy, K. and Pierce, B. (1993), "Wage inequality and the rise in returns to skill", *Journal of Political Economy* Vol. 101 pp. 410-442.

Ham, J. C. (1982) "Estimation of a Labour Supply Model with Censoring Due to Unemployment and Underemployment". *The Review of Economic Studies*, Vol.49, N.3, pp. 335-354.

Hanson, G. H. and Harrison, A. (1999) "Trade Liberalization and Wage Inequality in Mexico". *Industrial and Labor Relations Review*, 52: 271-288.

Haguenauer, L. Markwald, R. Pourchet, H. (1998) "Estimativa do Valor da Producao Industrial e Elaboracao de Coeficientes de Exportacao e Importacao da Industria Brasileira (1985-96)". IPEA, Texto para Discussao #563.

Heckman, James (1979) "Sample Selection Bias as a Specification Error". *Econometrica*, Vol.47, No.1. pp. 153-161.

Heckman, James J. and Sedlacek, Guilherme, (1985) "Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self Selection in the Labor Market" *Journal of Political Economy*, Dec, 93: 1077-1125.

Heckman, James J., and Hotz, Joseph V., (1988) "An Investigation of the Labor Market Earnings of Panamanian Males: Evaluating the Sources of Inequality", *Journal of Human Resources*, Fall, 21: 507-542.

Heckman, James J., and Honore, Bo., (1990) "The Empirical Content of the Roy Model ", *Econometrica*, Vol.58 N.5 pp. 1121-1149.

Heckman, James J. (1990) "Varieties of Selection Bias" *The American Economic Review.* Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association. Vol.80, N.2. pp. 313-318.

Heckman, J., Ichimura, H., and Todd, P. (1997) "Matching as an Econometric Evaluation Estimator", *Review of Economic Studies*, Vol. 64, pp. 605-654.

Heywood, John S. and Mohanty, Madhu S. (1995) "Estimation of the US Federal Job Queue in the Presence of an Endogenous Union Queue". *Economica*, Vol. 62, pp.479 - 493.

Holtzer H. J., Katz, L. F and Krueger, A. B., (1991) "Job Queues and Wages". *Quarterly Journal of Economics.* Pp. 739-768.

ILO (International Labour Office)(1972) "Employment, Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya". Geneva: ILO publications.

Kramarz, R. and Phillipon, T. (2000) "The Impact of Differential Payroll Tax

Subsidies on Minimum Wage Employment". Mimeo.

Krueger, Alan and Summers, Lawrence. (1987) "Reflections on the Inter-industry Wage Structure", In Kevin Lang and Jonathan S. Leonard, eds, *Unemployment and the Structure of Labor Markets*, Oxford: Basil Blackwell.

Kume, H., Piani, G. and Souza, C.F. (2000). "Instrumentos de Politica Comercial no Periodo 1987-1998". Mimeo. IPEA.

Lang, Kevin (1998) "The Effect of Trade Liberalization on Wages and Employment: The Case of New Zealand. *Journal of Labor Economics*, Vol. 16 No.4 pp.792-814.

Lanot, G. and Walker, I. (1998) "The union/non-union wage differential: An application of semi-parametric methods". *Journal of Econometrics* Vol.84 pp.327-349.

Lemos, S. (2002) "Robust Quantification of Minimum Wage Effects On Wages And Employment Using a New Data Set - A Menu Of Minimum Wage Variables". UCL. Mimeo.

Machin, S., Manning, A., and Rahman, L. (2002) "Where the Minimum Wage Bites Hard: The Introduction of the UK National Minimum Wage to a Low Wage Sector". CEP. Mimeo.

Maddala, G. S. (1983) *Limited Dependent and Qualitative Variables in Econonometrics.* Cambridge University Press.

Magnac, Thierry, (1991) "Segmented or Competitive Labor Market? *Econometrica*, Vol.59 N.1 pp. 165-187.

Maloney, W. F. (1997) "The Structure of Labor Market in Developing Countries: Time Series Evidence on Competing Views". Mimeo. World Bank.

Maloney, W. F. (1999) "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico". *The World Bank Economic Review* Vol. 13, N.2, pp. 275-312.

Marcoullier, D. et al. (1997) "Formal Measures of the Informal Sector Wage Gap in Mexico, El Salvador and Peru". *Economic Development and Cultural Change*, Vol. 45 No.2. pp. 367-392.

Marquez, G., Pages, C. (1997) "Trade and Employment: Evidence from Latin America and the Caribbean". Mimeo. Inter-American Development Bank.

Martins, M. F. O. (2001) "Parametric and Semiparametric Estimation of Sample Selection Models: An Empirical Application to the Female Labour Force in Portugal". *Journal of Applied Econometrics.* Vol.16, pp. 23-39.

Menezes-Filho, N. and Arbache, J. S.(2002) "Trade Liberalization, Product Markets and Labor Markets in Brazil". Mimeo.

Mengistae, Taye (1998) "Wage Rates and Job Queues: Does the Public Sector Overpay in Ethiopia?" St. Antony's College and CSAE, University of Oxford. *Working Paper Series/98-20.*

Mincer, J. (1976) "Unemployment Effects of Minimum Wages" *Journal of Political Economy,* Vol.84 pp.87-105.

Mohanty, Madhu S. (1998)"The Role of the Desire for Union Status in the Decision to Enter Local Government Job Queues: The US Evidence., *Applied Economics*Vol.30, pp.1369-1378.

Moreira, M.M. e Correa, P.G. (1998), 'A First Look at the Impacts of Trade Liberalization on Brazilian Manufacturing Industry', *World Development,* 26: 1859-1874.

Muendler, Marc-Andreas (2001) "Trade, Technology and Productivity: A Study of Brazilian Manufacturers, 1986-1998". University of California, Berkeley, mimeo.

Neri, Marcelo. (1997) A Efetividade do Salario Minimo no Brasil: Pobreza, Efeito-farol e Padroes Regionais. Mimeo. IPEA.

Newey, W, Powell, J. and Walker, J. (1990) "Semiparametric Estimation of Selection Models: Some Empirical Results", textitAmerican Economic Review, Papers and Proceedings, 80(2), pp.324-328.

Nickell, Steve. (1999) Product markets and labour markets. *Labour Economics.* Vol.6 pp.1-20.

Pavcnik, N., Bloom, A., Goldberg, P. and Schady, N. (2002). "Trade Liberalization and Labour Market Adjustment in Brazil", Dartmouth College, mimeo.

Pero, Valeria and Urani, Andre (1994) "Determinantes do Excesso de Mao-de-Obra do Setor Formal do Mercado de Trabalho Metropolitano" In *Perspectivas da Economia Brasileira,* Vol.2, pp. 541-559. IPEA.

Pratap, S. and Quintin E, (2001) "Are labor markets segmented in Argentina? A semiparametric approach," Center for Latin America Working Papers 0701, Federal Reserve Bank of Dallas.

Oliveira-Junior, M. (2000) "A liberalizacao Comercial Brasileira e os Coeficientes de Importacao – 1990/95". IPEA, Texto para Distrucao #703.

Pereira, S. (2001) "The impact of Minimum Wages on Youth Employment in Portugal". UCL, Mimeo.

Pissarides, C.A. (1997), 'Learning by trading and returns to human capital in developing countries', *World Bank Economic Review,* 11: 17-32.

Poirier, D. J. (1980) "Partial observability in bivariate probit models. *Journal of Econometrics,* Vol.12, pp. 209-17.

Portugal, P. and Cardoso, A. R., (2001) "Disentangling the Minimum Wage

Puzzle: An Analysis of Job Accessions and Separations from a Longitudinal Matched Employer-Employee Data Set". *CEPR Discussion paper series* N. 2844.

Pratap, S and Quintin E., (2001) "Are labor markets segmented in Argentina? A semiparametric approach" Center for Latin America Working Papers 0701, Federal Reserve Bank of Dallas.

Ramos, Lauro (2002) "A Evolucao da Informalidade no Brasil Metropolitano: 1991-2001". Texto para Discussao IPEA N.914.

Rauch, J. E. (1991) "Modelling the Informal Sector Formally". *Journal of Development Economics.* Vol.35 pp. 33-47.

Rebitzer, J. B. and Taylor, L. J. (1991) "The Consequences of Minimum Wage Laws: Some New Theoretical Ideas". *NBER Working paper* N.3877.

Revenga, Ana. (1992), "Exporting Jobs? The Impact of Import Competition on Employment and Wages in U.S. Manufacturing". *Quarterly Journal of Economics,* Vol. 107, pp. 255-284.

Revenga, Ana. (1997) Employment and Wage Effects of Trade Liberalization: The Case of Mexican Manufacturing. *Journal of Labor Economics.* Vol.17 (3) pp.S20-S43.

Robbins, D.J. and Gindling, T.H. (1999), 'Trade Liberalisation and the Relative Wages for more-Skilled workers in Costa Rica', *Review of Development Economics,* Vol.3 pp. 140-154.

Robbins, D. J. (1996a) "Stolper-Samuelson in the Tropics? Trade Liberalization and Wages in Colombia: 1976-1994". *Development Discussion Paper No. 557,* Harvard Institute for International Economics.

Robins, D. J. (1996b) "HOS Hits Facts: Facts Win; Evidence on Trade and Wages in the Developing Countries". *Development Discussion Paper No. 557,* Harvard Institute for International Economics.

Robbins, D.J. and Gindling, T.H. (1999), 'Trade Liberalisation and the Relative Wages for more-Skilled workers in Costa Rica', *Review of Development Economics,* 3: 140-154.

Rosenbaum, P., and D. Rubin, (1983), "The central role of the propensity score in observational studies for causal effects", *Biometrika,* Vol. 70 N.1 pp.41-55

Rossi Jr, Jose L. and Ferreira, Pedro C. (1999) "Evolucao da Produtividade Industrial Brasileira e Abertura Comercial" *Pesquisa e Planejamento Economico.* V. 29, No.1, pp. 1-36.

Roy, A. D. (1951) "Some Thought on the Distribution of Earnings" *Oxford Economic Papers (New Series)* Vol.3 N.2 pp. 135-146.

Saavedra, J. and Chong, A. (1999), "Structural reform, institutions and earnings: Evidence from the formal and informal sectors in urban Peru", *Journal of Development Studies*, Vol. 35 N.4 pp.95-116.

Saint-Paul, Gilles (1997) "Dual Labor Markets: A Macroeconomic Perspective". The MIT Press.

Saxonhouse, Gary R. (1976) "Estimated Parameters as Dependent Variables". *American Economic Review*, Vol.66 N.1 pp. 178-183.

Schafgans, M. M. A. (1998) "Ethnic Wage Differences in Malaysia: Parametric and Semiparametric Estimation of the Chinese-Malay Wage Gap". *Journal of Applied Econometrics*. Vol. 13, pp.481-504.

Slaughter, Matthew, J. (2001) International Trade and Labor-Demand Elasticities. *Journal of International Economics*. Vol.54 (11) pp.27-56.

Smith, J. and Todd, P. (2000) "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?". Mimeo.

Soares, Sergei . S. D. (2002b) "O Impacto Distributivo do Salario Minimo: A Distribuicao Individual dos Rendimentos do Trabalho". Texto para Discussao N. 873. IPEA.

Stewart, M. B. (2002) "The Impact of the Introduction of the UK Minimum Wage on the Employment Probabilities of Low Wage Workers". *Warwick Economic Research Papers* N.630

Tannuri-Pianto, M. and Pianto, D.M. (2002) "Informal Employment in Brazil - a choice at the top and segmentation at the bottom: A Quantile Regression Approach. Mimeo.

Tunali, Insan (1986) "A General Structure for Models of Double-Selection and an Application to Joint Migration/Earning Process with Remigration". In R. Ehrenberg (ed.), *Research in Labor Economics*. Vol. 8, Part B, pp. 235-282. Greenwich, MA.: JAI Press.

Venti, S. F. (1987) "Wages in the federal and private sectors, in Public Sector Payrolls". In *Public Sector Payrolls*, ed. D. A. Wise, University of Chicago Press, Chicago, pp. 147-182.

Verry, Donald and Araujo, Tarcisio (1996) "Dualism in Brazilian Metropolitan Labour Markets. *UCL Discussion Paper* 96-15.

Welch, Finis (1974) "Minimum Wage Legislation in the United States". *Economic Inquiry*, Vol. 12, pp.285-318.

Wood (1997) "Openness and Wage Inequality in Developing Countries: The Latin American Challenge to East Asian Conventional Wisdom". *World Bank Economic Review* 11: 33-58.

World Bank (2002) World Development Report.

Yamada, Gustavo (1996) "Urban Informal Employment and Self-Employment in Developing Countries: Theory and Evidence". *Economic Development and Cultural Change.* Vol. 44 No. 2 , pp. 289-314.

Zavodny, M. (2000) "The Effect of the Minimum Wage on Employment and Hours" St. *Labor Economics.* Vol.7, N.6 pp.729-750.

Appendix A

Appendix - Chapter 2

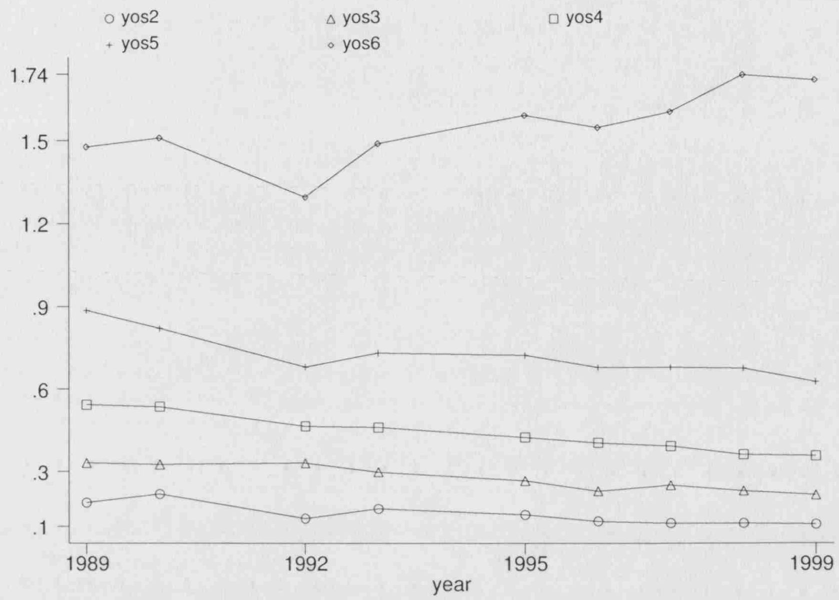

Figure A.1: Relative Returns to Education (Survey based: Full Sample) - 1989 - 1999

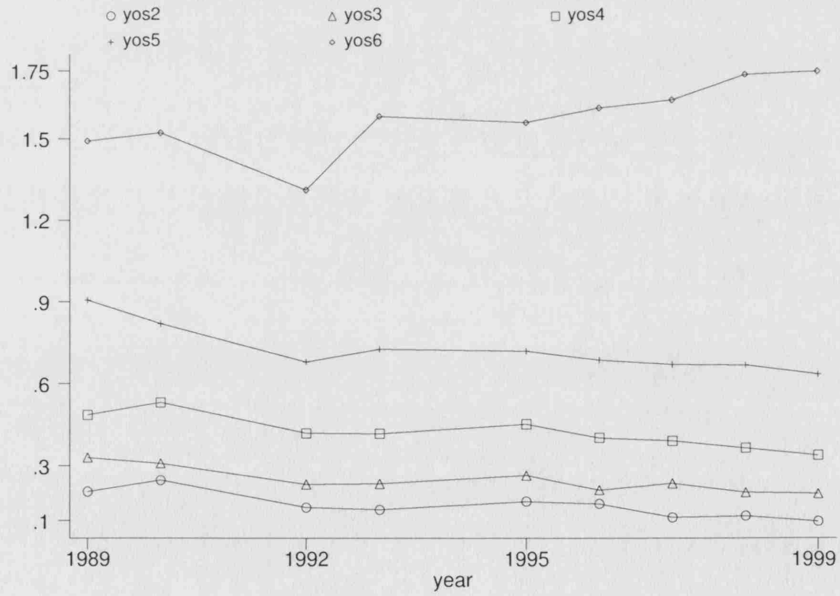

Figure A.2: Relative Returns to Education (Survey based: Registered Sample) - 1989 - 1999

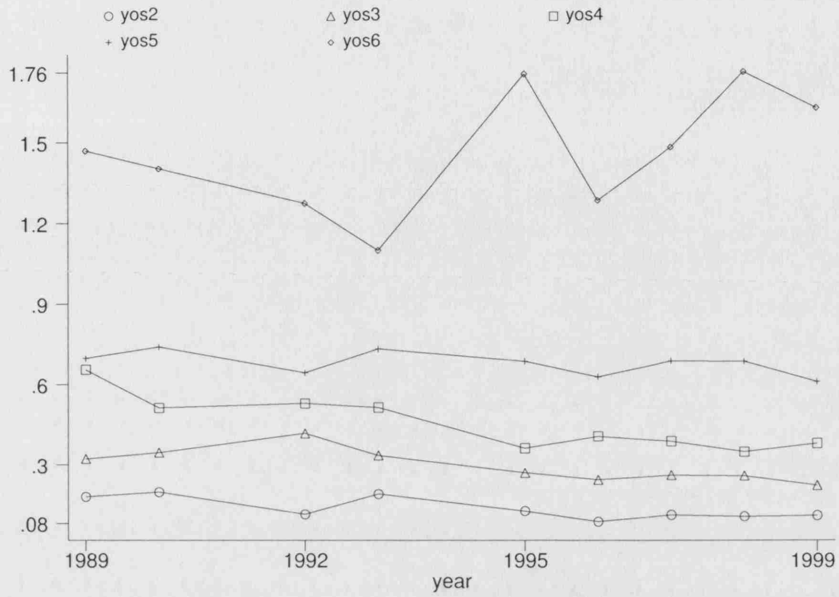

Figure A.3: Relative Returns to Education (Survey based: Non-Registered Sample) - 1989 - 1999

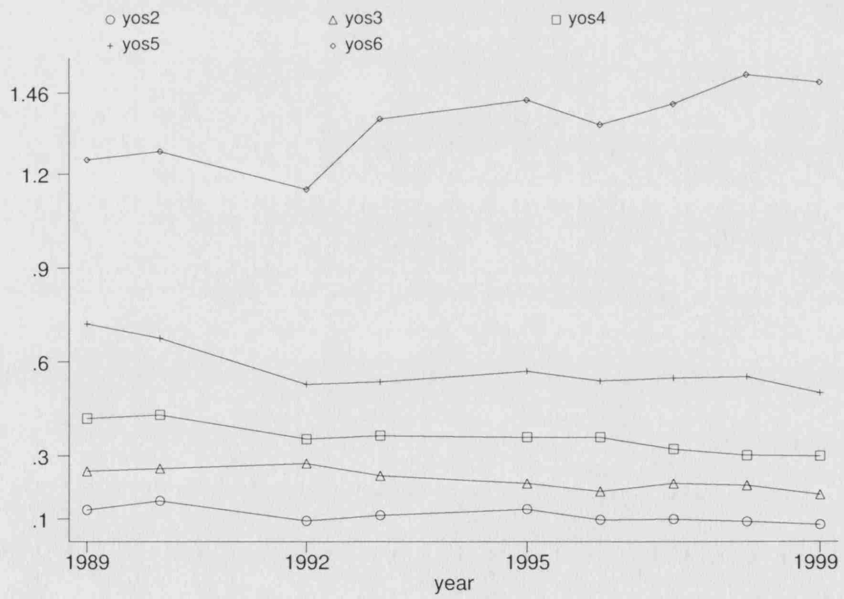

Figure A.4: Relative Returns to Education (Survey based and more controls: Full Sample) - 1989 - 1999

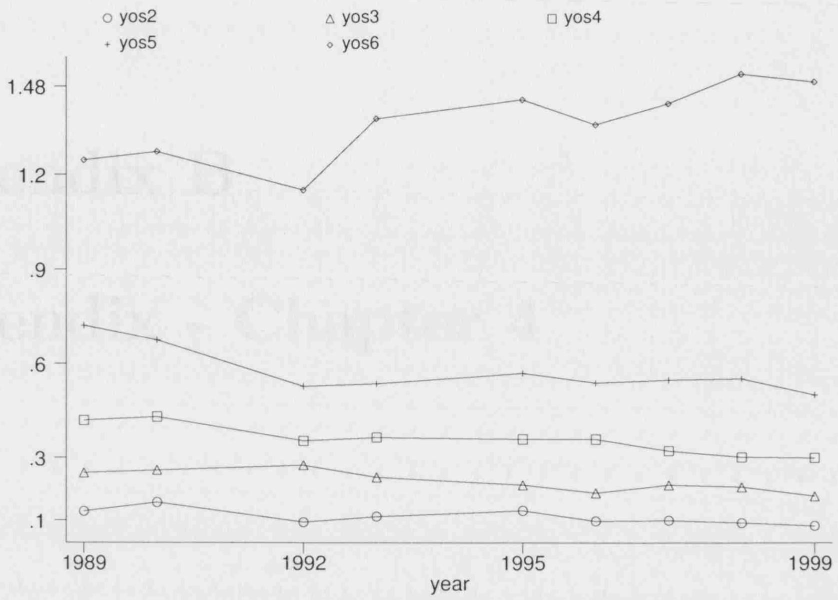

Figure A.5: Relative Returns to Education (Survey based and more controls: Registered Sample) - 1989 - 1999

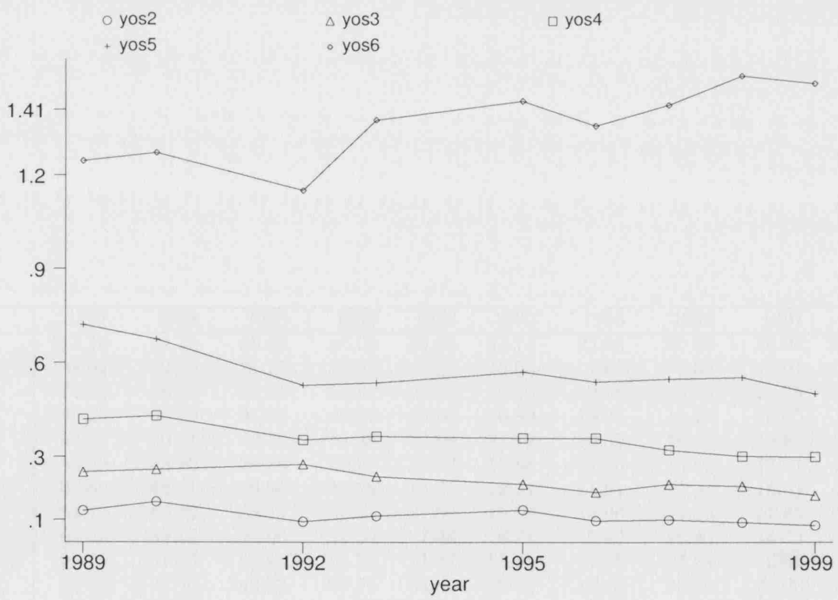

Figure A.6: Relative Returns to Education (Survey based and more controls: Non-Registered Sample) - 1989 - 1999

Appendix B

Appendix - Chapter 4

industry	1987	1988	1989	1990	1992	1993	1995	1996	1997	1998
apparel	117.19	94.32	95.46	67.00	36.60	23.74	23.06	23.06	26.07	26.07
chemical	34.02	39.88	31.26	27.91	14.77	15.06	7.62	7.62	14.01	15.89
electronics	71.00	56.08	48.66	52.87	29.71	24.10	19.38	19.38	21.54	21.02
food	89.28	58.02	51.87	52.22	20.06	16.89	15.76	15.76	18.65	18.94
mechanic	47.52	50.21	43.95	41.49	22.09	21.72	16.69	16.69	18.65	18.63
metal	54.59	43.80	29.18	29.28	17.07	13.88	14.85	14.85	18.01	18.02
nonmetal	81.69	46.21	39.56	38.79	13.24	12.21	11.91	11.91	15.51	15.43
others	64.79	63.99	58.17	58.90	27.88	19.08	14.96	14.96	17.85	17.87
paper	65.47	30.12	23.05	22.57	7.96	8.21	10.40	10.40	14.73	14.73
pharmaceutical	91.70	51.83	39.80	35.83	14.77	13.56	7.33	7.33	9.96	9.96
plastics	31.40	72.10	49.53	50.75	24.21	20.22	19.10	19.10	21.90	21.93
rubber	122.41	58.52	67.08	70.24	26.04	16.94	13.95	13.95	16.35	16.01
shoes	96.94	39.76	38.49	28.81	16.47	15.01	18.24	18.24	20.80	19.43
textiles	123.08	83.91	85.70	49.23	31.44	21.34	21.26	21.26	24.38	24.37
transport	198.91	128.09	151.65	208.59	61.61	50.84	124.94	124.94	104.40	78.67
wood	53.14	28.89	29.14	29.41	9.49	9.82	11.87	11.87	15.14	15.13

Table B.1: Effective Tariff in %

industry	1987	1988	1989	1990	1992	1993	1995	1996	1997	1998
apparel	102.73	75.99	74.97	51.12	29.29	20.00	19.75	19.75	22.75	22.75
chemical	38.93	33.53	26.07	22.88	12.70	12.28	7.23	7.23	12.59	14.00
electronics	59.45	49.26	40.25	42.23	23.92	19.81	16.35	16.35	18.76	18.43
food	62.66	37.67	30.65	30.42	15.40	13.62	12.89	12.89	15.68	15.94
mechanic	48.97	46.83	38.83	37.23	20.16	19.14	15.56	15.56	17.77	17.73
metal	43.28	36.01	23.42	23.25	12.41	10.34	11.24	11.24	13.99	13.97
nonmetal	63.75	39.25	32.27	31.49	11.81	10.69	10.53	10.53	13.71	13.64
others	53.15	49.09	42.07	41.59	21.07	16.44	13.46	13.46	16.32	16.35
paper	59.46	32.08	24.28	23.62	9.49	9.29	10.29	10.29	14.18	14.18
pharmaceutical	72.32	45.30	34.43	31.45	13.82	12.78	7.95	7.95	10.71	10.75
plastics	36.63	57.13	39.50	38.97	19.24	16.82	15.24	15.24	18.10	18.16
rubber	82.00	49.32	47.62	46.55	20.62	14.90	12.45	12.45	15.00	14.84
shoes	74.14	40.98	35.83	29.62	15.98	14.18	15.27	15.27	17.98	17.21
textiles	87.39	57.33	53.30	31.82	20.94	15.63	16.15	16.15	19.16	19.17
transport	78.23	54.68	52.42	59.50	30.51	26.51	35.53	35.53	33.86	28.99
wood	50.00	30.32	25.82	25.43	9.77	9.46	10.97	10.97	13.97	13.97

Table B.2: Nominal Tariff in %

industry	1987	1988	1989	1990	1992	1993	1995	1996	1997	1998
apparel	0.22	0.20	0.21	0.36	0.50	0.53	3.25	3.88	5.02	5.03
chemical	11.77	10.77	11.42	10.19	14.00	15.74	19.26	20.01	21.16	22.86
electronics	11.09	10.80	10.47	8.62	19.82	24.26	34.01	39.06	55.51	68.75
food	2.51	1.24	2.97	2.44	2.78	3.30	4.62	3.94	4.53	4.80
mechanic	8.40	8.37	6.28	8.66	14.51	13.02	21.63	29.07	33.87	36.80
metal	2.88	2.34	2.88	2.53	3.30	3.18	4.83	4.86	6.32	7.31
nonmetal	1.05	0.79	0.92	0.89	1.14	1.23	2.02	2.48	2.56	2.20
others	11.45	11.42	13.69	12.73	18.59	16.78	25.80	28.68	27.78	29.11
paper	2.94	2.24	2.34	2.13	2.17	2.45	5.96	6.74	7.74	8.27
pharmaceutical	4.64	3.82	4.31	3.93	4.48	4.81	7.22	8.84	9.13	9.78
plastics	0.92	0.70	0.55	0.75	1.52	1.82	3.22	4.00	4.75	5.10
rubber	4.32	4.94	5.92	4.90	4.81	4.75	9.13	8.95	10.42	12.12
shoes	5.21	3.98	5.77	4.36	6.03	6.04	10.11	9.91	11.45	10.37
textiles	1.10	1.52	2.40	2.32	4.40	9.58	13.84	15.40	17.34	16.07
transport	8.67	4.43	4.20	3.99	7.17	8.39	14.61	12.56	16.58	22.84
wood	0.46	0.32	0.45	0.38	0.47	0.60	0.87	1.10	1.54	1.78

Table B.3: Import Penetration Ratio in %

industry	1987	1988	1989	1990	1992	1993	1995	1996	1997	1998
apparel	3.14	3.31	1.21	0.99	2.26	2.44	1.93	1.99	2.15	1.95
chemical	3.82	4.50	5.30	4.16	6.56	6.78	6.66	7.62	7.79	7.90
electronics	9.70	9.24	8.26	5.44	14.88	14.75	10.63	11.19	14.37	18.47
food	20.89	21.24	15.68	12.09	14.82	14.43	13.79	13.81	14.89	14.26
mechanic	5.44	6.32	5.05	4.03	9.20	9.01	8.20	9.98	10.99	11.29
metal	11.36	18.97	16.23	13.63	19.97	17.99	14.53	14.56	13.58	13.60
nonmetal	2.07	2.29	1.89	1.41	2.26	2.76	2.27	2.21	2.15	2.11
others	13.06	8.64	8.82	7.99	14.77	12.76	9.15	9.11	9.47	11.08
paper	8.59	11.17	8.16	7.17	10.50	9.89	12.35	10.02	11.22	11.97
pharmaceutical	2.26	1.79	1.41	0.97	1.76	1.84	1.71	1.74	1.70	1.93
plastics	2.68	2.61	0.86	0.52	2.04	2.23	1.40	1.27	1.65	1.85
rubber	7.04	7.42	8.35	7.18	10.71	9.32	7.98	8.67	9.21	10.62
shoes	35.17	37.18	26.72	29.63	61.01	55.90	46.51	56.52	63.80	69.67
textiles	9.33	8.04	7.52	6.68	13.37	11.03	9.96	9.63	10.72	11.07
transport	20.32	15.33	14.08	9.73	16.66	13.45	8.94	10.60	13.87	20.01
wood	6.55	8.37	3.61	3.11	7.51	8.74	7.36	7.24	8.19	8.70

Table B.4: Export Orientation Ratio in %

industry	1987	1988	1989	1990	1992	1993	1995	1996	1997	1998
apparel	0.31	0.38	0.18	0.08	0.27	0.26	0.24	0.13	0.12	0.20
chemical	0.52	0.51	0.46	0.54	0.45	0.47	0.15	0.16	0.28	0.26
electronics	0.38	0.43	0.23	0.42	0.44	0.29	0.17	0.27	0.22	0.06
food	0.42	0.48	0.39	0.34	0.50	0.38	0.31	0.23	0.34	0.28
mechanic	0.31	0.61	0.24	0.37	0.60	0.34	0.23	0.29	0.21	0.22
metal	0.46	0.55	0.36	0.19	0.56	0.49	0.26	0.26	0.29	0.36
nonmetal	0.31	0.35	0.31	0.26	0.63	0.46	0.21	0.17	0.23	0.14
others	0.20	0.47	0.41	0.36	0.51	0.47	0.33	0.21	0.36	0.35
paper	0.28	0.19	0.32	0.25	0.26	0.38	0.28	0.11	0.19	0.20
pharmaceutical	0.19	0.66	0.30	0.28	0.52	0.36	0.50	0.07	0.37	-0.13
plastics	0.18	0.51	0.00	0.25	0.31	0.19	0.12	-0.02	0.22	0.22
rubber	0.02	0.24	0.29	0.09	0.53	0.55	0.39	0.04	0.47	0.20
shoes	0.49	0.61	0.49	0.24	0.53	0.60	0.23	0.24	0.26	0.26
textiles	0.32	0.58	0.49	0.55	0.70	0.35	0.37	0.29	0.32	0.34
transport	0.46	0.52	0.43	0.32	0.61	0.41	0.31	0.35	0.18	0.23
wood	0.19	0.21	0.15	0.13	0.30	0.30	0.18	0.11	0.20	0.14

Table B.5: Dummy Coefficient for Registered Workers

industry	1987	1988	1989	1990	1992	1993	1995	1996	1997	1998
apparel	0.75	0.74	0.75	0.71	0.74	0.73	0.73	0.68	0.70	0.71
chemical	0.92	0.90	0.93	0.94	0.95	0.92	0.93	0.92	0.89	0.91
electronics	0.96	0.94	0.95	0.97	0.91	0.93	0.92	0.88	0.87	0.90
food	0.82	0.81	0.80	0.79	0.80	0.78	0.79	0.76	0.77	0.76
mechanic	0.95	0.95	0.97	0.94	0.91	0.94	0.90	0.85	0.89	0.88
metal	0.90	0.92	0.91	0.89	0.88	0.87	0.85	0.84	0.81	0.81
nonmetal	0.65	0.66	0.65	0.63	0.63	0.59	0.64	0.66	0.65	0.67
others	0.79	0.80	0.82	0.77	0.68	0.73	0.81	0.71	0.75	0.73
paper	0.90	0.89	0.92	0.91	0.88	0.85	0.85	0.82	0.81	0.78
pharmaceutical	0.89	0.91	0.86	0.92	0.91	0.91	0.88	0.87	0.87	0.89
plastics	0.94	0.94	0.94	0.92	0.90	0.91	0.90	0.87	0.89	0.86
rubber	0.95	0.92	0.98	0.95	0.88	0.94	0.94	0.95	0.85	0.88
shoes	0.81	0.83	0.82	0.85	0.81	0.80	0.81	0.79	0.81	0.79
textiles	0.93	0.88	0.91	0.91	0.90	0.86	0.90	0.87	0.89	0.86
transport	0.96	0.96	0.96	0.96	0.94	0.96	0.94	0.95	0.93	0.92
wood	0.72	0.68	0.69	0.69	0.65	0.60	0.59	0.59	0.58	0.62

Table B.6: Proportion of Registered Workers

Appendix C

Appendix - Chapter 5

		Pooled Sample			Formal			Informal		
	Transitions to:		t^*m	z		t^*m	z		t^*m	z
1982	Noemployment	May	-0.006	[0.67]	May	-0.003	[0.32]	May	-0.023	[0.99]
	Unemployment	(N=14359)	-0.002	[0.29]	(N=11148)	-0.002	[0.31]	(N=3211)	-0.006	[0.43]
	Inactivity		-0.003	[0.52]		-0.001	[0.12]		-0.015	[0.86]
	Nonemployment	Nov	0	[0.01]	Nov	0.004	[0.54]	Nov	-0.014	[0.67]
	Unemployment	(N=18802)	-0.006	[1.37]	(N=15051)	-0.003	[0.56]	(N=3751)	-0.019	[1.66]
	Inactivity		0.006	[1.18]		0.006	[1.20]		0.008	[0.48]
1983	Noemployment	May	0.008	[0.98]	May	-0.001	[0.11]	May	0.043	[1.64]
	Unemployment	(N=16437)	0.005	[1.02]	(N=13186)	0.001	[0.15]	(N=3251)	0.024	[1.62]
	Inactivity		0.003	[0.47]		-0.001	[0.21]		0.02	[1.04]
	Nonemployment	Nov	0.016	[1.54]	Nov	0.007	[0.76]	Nov	0.044	[1.42]
	Unemployment	(N=17539)	0.001	[0.10]	(N=13871)	0.001	[0.11]	(N=3580)	0.003	[0.20]
	Inactivity		0.017	[2.08]*		0.007	[0.98]		0.047	[1.80]
1984	Noemployment	May	-0.023	[1.98]*	May	-0.005	[0.42]	May	-0.08	[2.55]*
	Unemployment	(N=16650)	-0.008	[1.16]	(N=13529)	0.004	[0.46]	(N=3391)	-0.05	[2.74]**
	Inactivity		-0.012	[1.37]		-0.009	[1.01]		-0.018	[0.78]
	Nonemployment	Nov	0.015	[1.24]	Nov	0.012	[1.04]	Nov	0.021	[0.48]
	Unemployment	(N=17187)	0.005	[0.73]	(N=13621)	0.007	[0.95]	(N=3566)	-0.008	[0.36]
	Inactivity		0.008	[0.85]		0.003	[0.36]		0.051	[1.30]
1985	Noemployment	May	0.005	[0.73]	May	-0.005	[0.69]	May	0.077	[2.96]**
	Unemployment	(N=23924)	0	[0.05]	(N=19532)	-0.004	[0.85]	(N=4392)	0.025	[1.61]
	Inactivity		0.006	[1.06]		-0.001	[0.13]		0.049	[2.39]*
	Nonemployment	Nov	0.005	[0.47]	Nov	0.009	[0.87]	Nov	-0.013	[0.44]
	Unemployment	(N=17514)	0	[0.08]	(N=13917)	0.003	[0.55]	(N=3597)	-0.012	[0.96]
	Inactivity		0.005	[0.55]		0.006	[0.70]		-0.002	[0.08]
1995	Noemployment	May	-0.01	[0.64]	May	-0.006	[0.40]	May	-0.017	[0.58]
	Unemployment	(N=8588)	-0.004	[0.69]	(N=5062)	-0.003	[0.49]	(N=3526)	0	[0.00]
	Inactivity		-0.005	[0.33]		0.001	[0.08]		-0.02	[0.74]
	Nonemployment	Oct	0.017	[1.05]	Oct	0.017	[0.96]	Oct	0.026	[0.84]
	Unemployment	(N=8425)	0.006	[0.79]	(N=4941)	0.005	[0.68]	(N=3467)	0.007	[0.67]
	Inactivity		0.011	[0.72]		0.012	[0.81]		0.017	[0.59]
1996	Noemployment	May	0.013	[0.77]	May	0.008	[0.42]	May	0.029	[0.91]
	Unemployment	(N=6890)	0.008	[0.97]	(N=3964)	0.003	[0.60]	(N=2885)	0.017	[1.03]
	Inactivity		0.005	[0.33]		0.004	[0.24]		0.012	[0.45]
	Nonemployment	Oct	0.034	[2.14]*	Oct	-0.001	[0.10]	Oct	0.078	[2.60]**
	Unemployment	(N=8622)	0.021	[2.77]**	(N=5066)	0.002	[0.38]	(N=3556)	0.039	[2.70]**
	Inactivity		0.014	[1.03]		-0.006	[0.48]		0.041	[1.56]
1997	Noemployment	May	0.023	[1.47]	May	0.015	[0.95]	May	0.028	[0.94]
	Unemployment	(N=9664)	-0.002	[0.31]	(N=5917)	0.008	[1.75]	(N=3747)	-0.013	[1.33]
	Inactivity		0.027	[1.89]		0.005	[0.41]		0.05	[1.84]
	Nonemployment	Oct	0.017	[1.07]	Oct	0.003	[0.18]	Oct	0.039	[1.35]
	Unemployment	(N=9428)	-0.001	[0.16]	(N=5542)	0.003	[0.41]	(N=3855)	-0.005	[0.46]
	Inactivity		0.017	[1.26]		-0.001	[0.13]		0.047	[1.79]
1998	Noemployment	May	0.012	[0.73]	May	-0.003	[0.20]	May	0.024	[0.82]
	Unemployment	(N=8489)	0.01	[1.14]	(N=4812)	0	[0.06]	(N=3677)	0.018	[1.13]
	Inactivity		0.001	[0.10]		-0.005	[0.36]		0.006	[0.23]
	Nonemployment	Oct	0.034	[2.07]*	Oct	-0.005	[0.33]	Oct	0.076	[2.53]*
	Unemployment	(N=9551)	0.017	[1.80]	(N=5546)	0.004	[0.56]	(N=3953)	0.027	[1.72]
	Inactivity		0.016	[1.21]		-0.008	[0.68]		0.047	[1.84]
1999	Noemployment	May	-0.005	[0.31]	May	0.021	[1.17]	May	-0.029	[1.09]
	Unemployment	(N=9681)	-0.005	[0.76]	(N=5759)	0.002	[0.23]	(N=3922)	-0.012	[1.14]
	Inactivity		0.001	[0.06]		0.02	[1.30]		-0.014	[0.60]
	Nonemployment	Oct	0	[0.02]	Oct	0.004	[0.27]	Oct	-0.007	[0.28]
	Unemployment	(N=10255)	0.002	[0.32]	(N=6230)	0.007	[0.93]	(N=4025)	-0.002	[0.18]
	Inactivity		-0.003	[0.21]		-0.002	[0.13]		-0.005	[0.21]

Table C.1: Changes in Transition to Non-employment: Probit (Excluding 1.5, 2 and 3 minimum wages)

		Pooled Sample			Formal Sector			Informal Sector		
	Transitions to:		t^*m	z		t^*m	z		t^*m	z
1982	Noemployment	May	-0.002	[0.28]	May	-0.001	[0.11]	May	-0.014	[0.66]
	Unemployment	(N=16962)	0	[0.07]	(N=13461)	-0.001	[0.18]	(N=3501)	-0.003	[0.23]
	Inactivity		-0.001	[0.17]		0	[0.04]		-0.009	[0.56]
	Nonemployment	Nov	-0.001	[0.14]	Nov	0.002	[0.37]	Nov	-0.017	[0.73]
	Unemployment	(N=19303)	-0.006	[1.42]	(N=15972)	-0.003	[0.72]	(N=3331)	-0.019	[1.46]
	Inactivity		0.005	[1.07]		0.005	[1.13]		0.005	[0.27]
1983	Noemployment	May	0.005	[0.63]	May	-0.002	[0.23]	May	0.019	[0.72]
	Unemployment	(N=18304)	0.003	[0.66]	(N=15077)	0	[0.02]	(N=3227)	0.013	[0.91]
	Inactivity		0.002	[0.35]		-0.001	[0.22]		0.009	[0.47]
	Nonemployment	Nov	0.011	[1.12]	Nov	0.004	[0.45]	Nov	0.032	[0.94]
	Unemployment	(N=18070)	-0.001	[0.16]	(N=14859)	-0.001	[0.09]	(N=3211)	-0.006	[0.33]
	Inactivity		0.014	[1.73]		0.005	[0.73]		0.047	[1.65]
1984	Noemployment	May	-0.019	[1.73]	May	-0.003	[0.23]	May	-0.074	[2.09]*
	Unemployment	(N=17456)	-0.006	[0.94]	(N=14389)	0.004	[0.60]	(N=3067)	-0.048	[2.32]*
	Inactivity		-0.011	[1.32]		-0.008	[0.94]		-0.014	[0.55]
	Nonemployment	Nov	0.011	[0.98]	Nov	0.01	[0.96]	Nov	-0.003	[0.06]
	Unemployment	(N=18072)	0.004	[0.62]	(N=14878)	0.006	[0.99]	(N=3194)	-0.018	[0.69]
	Inactivity		0.005	[0.62]		0.002	[0.25]		0.04	[0.91]
1985	Noemployment	May	0.001	[0.12]	May	-0.008	[1.22]	May	0.068	[2.30]*
	Unemployment	(N=24828)	-0.002	[0.35]	(N=20951)	-0.005	[1.10]	(N=3877)	0.021	[1.17]
	Inactivity		0.002	[0.48]		-0.003	[0.67]		0.043	[1.92]
	Nonemployment	Nov	-0.006	[0.63]	Nov	-0.002	[0.21]	Nov	-0.021	[0.65]
	Unemployment	(N=18543)	-0.003	[0.75]	(N=15406)	-0.001	[0.16]	(N=3137)	-0.014	[1.02]
	Inactivity		-0.002	[0.3]		-0.001	[0.14]		-0.007	[0.23]
1995	Noemployment	May	-0.013	[1.19]	May	-0.008	[0.71]	May	-0.014	[0.58]
	Unemployment	(N=12299)	-0.005	[1.12]	(N=8125)	-0.003	[0.5]	(N=4174)	-0.007	[0.81]
	Inactivity		-0.007	[0.72]		-0.004	[0.47]		-0.008	[0.37]
	Nonemployment	Oct	0	[0.00]	Oct	0	[0.01]	Oct	0.002	[0.08]
	Unemployment	(N=13581)	-0.001	[0.30]	(N=8825)	0.002	[0.32]	(N=4756)	-0.006	[0.81]
	Inactivity		0.002	[0.17]		-0.002	[0.2]		0.009	[0.45]
1996	Noemployment	May	-0.005	[0.42]	May	-0.004	[0.33]	May	0	[0.02]
	Unemployment	(N=12712)	0.003	[0.46]	(N=7931)	0.004	[0.68]	(N=4713)	0.006	[0.52]
	Inactivity		-0.008	[0.80]		-0.007	[0.68]		-0.005	[0.26]
	Nonemployment	Oct	0.022	[1.69]	Oct	-0.002	[0.21]	Oct	0.054	[2.12]*
	Unemployment	(N=12712)	0.011	[1.74]	(N=7601)	0.001	[0.14]	(N=4585)	0.019	[1.66]
	Inactivity		0.011	[1.04]		-0.005	[0.54]		0.035	[1.60]
1997	Noemployment	May	0.003	[0.20]	May	0.001	[0.07]	May	0.007	[0.24]
	Unemployment	(N=11074)	-0.004	[0.74]	(N=7068)	0.011	[1.48]	(N=4006)	-0.014	[1.62]
	Inactivity		0.009	[0.73]		-0.005	[0.43]		0.029	[1.17]
	Nonemployment	Oct	-0.006	[0.46]	Oct	-0.006	[0.49]	Oct	0.001	[0.03]
	Unemployment	(N=11591)	-0.004	[0.71]	(N=7264)	0.001	[0.16]	(N=4327)	-0.01	[0.92]
	Inactivity		-0.001	[0.14]		-0.007	[0.73]		0.013	[0.57]
1998	Noemployment	May	0.008	[0.56]	May	-0.009	[0.69]	May	0.037	[1.35]
	Unemployment	(N=11286)	0.011	[1.35]	(N=6906)	0	[0.04]	(N=4380)	0.028	[1.76]
	Inactivity		-0.003	[0.29]		-0.009	[0.87]		0.008	[0.36]
	Nonemployment	Oct	0.027	[1.87]	Oct	-0.004	[0.33]	Oct	0.065	[2.36]*
	Unemployment	(N=11828)	0.018	[1.95]	(N=7365)	0.006	[0.84]	(N=4463)	0.026	[1.66]
	Inactivity		0.009	[0.81]		-0.009	[0.96]		0.037	[1.60]
1999	Noemployment	May	-0.021	[1.56]	May	-0.001	[0.03]	May	-0.036	[1.47]
	Unemployment	(N=12272)	-0.009	[1.49]	(N=7060)	-0.003	[0.41]	(N=4212)	-0.016	[1.56]
	Inactivity		-0.01	[0.89]		0.004	[0.3]		-0.017	[0.77]
	Nonemployment	Oct	-0.011	[0.90]	Oct	-0.005	[0.34]	Oct	-0.019	[0.80]
	Unemployment	(N=11575)	-0.001	[0.21]	(N=72640)	0.004	[0.65]	(N=4311)	-0.007	[0.57]
	Inactivity		-0.01	[0.90]		-0.008	[0.65]		-0.011	[0.55]

Table C.2: Changes in Transition to Non-employment: Probit (conditional on Wages two Months before the Minimum Wage Increase)

		Without Multiples of MW (1.5, 2, and 3 mw)			Conditonal on wages 2 months before the mw increase		
			t^*m	z		t^*m	z
1982	Informal Sector	May	0.009	[1.20]	May	0.011	[1.34]
	Self-Employment	(N=13461)	0.001	[0.64]	(N=11148)	-0.002	[2.85]**
	Informal Sector	Nov	-0.004	[0.72]	Nov	-0.007	[1.25]
	Self-Employment	(N=15972)	0.001	[0.29]	(N=15051)	0	[0.21]
1983	Informal Sector	May	0.01	[1.41]	May	-0.005	[0.72]
	Self-Employment	(N=15077)	0.001	[0.31]	(N=13186)	0	[0.49]
	Informal Sector	Nov	-0.008	[0.96]	Nov	0.006	[0.68]
	Self-Employment	(N=14859)	0.001	[0.29]	(N=13959)	0.002	[0.74]
1984	Informal Sector	May	0.014	[1.27]	May	0.002	[0.24]
	Self-Employment	(N=14389)	-0.004	[1.25]	(N=13259)	0.002	[0.95]
	Informal Sector	Nov	-0.017	[1.48]	Nov	-0.021	[2.01]*
	Self-Employment	(N=14878)	0.003	[1.08]	(N=13621)	-0.001	[0.36]
1985	Informal Sector	May	0.001	[0.07]	May	-0.005	[0.84]
	Self-Employment	(N=20951)	0.004	[1.16]	(N=19532)	-0.001	[0.28]
	Informal Sector	Nov	-0.019	[1.82]	Nov	-0.026	[2.81]**
	Self-Employment	(N=15406)	-0.001	[0.42]	(N=13917)	0	[0.13]
1995	Informal Sector	May	-0.002	[0.14]	May	-0.02	[1.05]
	Self-Employment	(N=8125)	0.003	[0.53]	(N=5062)	-0.004	[0.62]
	Informal Sector	Oct	0.018	[1.36]	Oct	0.002	[0.11]
	Self-Employment	(N=8825)	0.003	[0.51]	(N=4958)	-0.016	[3.14]**
1996	Informal Sector	May	0.021	[1.31]	May	-0.024	[1.26]
	Self-Employment	(N=7999)	0.013	[1.56]	(N=4019)	-0.003	[0.43]
	Informal Sector	Oct	-0.009	[0.72]	Oct	-0.004	[0.22]
	Self-Employment	(N=7601)	0.005	[0.71]	(N=5066)	0.033	[2.53]*
1997	Informal Sector	May	-0.025	[1.82]	May	0.006	[0.31]
	Self-Employment	(N=7068)	-0.009	[1.57]	(N=5917)	-0.001	[0.15]
	Informal Sector	Oct	-0.034	[2.53]*	Oct	-0.014	[0.76]
	Self-Employment	(N=7264)	-0.008	[1.28]	(N=5542)	-0.003	[0.39]
1998	Informal Sector	May	0.012	[0.71]	May	0.016	[0.72]
	Self-Employment	(N=6906)	0.014	[1.48]	(N=4812)	0.006	[0.75]
	Informal Sector	Oct	0.017	[1.02]	Oct	0.037	[1.79]
	Self-Employment	(N=7365)	-0.004	[0.71]	(N=5598)	0.012	[1.37]
1999	Informal Sector	May	0.013	[0.79]	May	-0.01	[0.53]
	Self-Employment	(N=7060)	0.003	[0.45]	(N=5759)	-0.009	[1.59]
	Informal Sector	Oct	0.021	[1.24]	Oct	-0.002	[0.12]
	Self-Employment	(N=7264)	-0.004	[0.87]	(N=6230)	-0.009	[2.74]**

Table C.3: Changes in Transition to Informality: Probit (Robustness Check - Formal Sector)